Support Document #80

Consistency of Agency approach for formulated products with the European Union approach for using formulated product effects data in *ecological risk assessment*

The proposed Agency approach can be compared with the use of formulated product toxicity data in the ecological risk assessment employed in the European Union (EU). As background, the EU has provisions in its regulations (EU Directive 91/414/EEC) for testing of formulated products for effects on terrestrial and aquatic wildlife. Annex III of the EU data requirements indicate that aquatic organism testing of formulated products is required, but may be limited to those aquatic organism taxonomic groups where available active ingredient testing indicate sensitivity. Requirements for the testing of birds with formulated product may be limited to situations where exposure estimates for active ingredient exceed acitive ingredient-based toxicity endpoints by factors 10 to 100 (based on acute oral dose and short term dietary tests, respectively). Personal communication with Mark A Clook, (Principal Scientist with the United Kingdom's Ecotoxicology Pesticides Safety Directorate, Department for Environment Food and Rural Affairs) indicates that it is rare that testing for formulated product is required for birds, but testing of aquatic organisms with formulated product is routine and that these data requirements are generally accepted throughout the European Union countries.

Dr. Clook indicated that the quantitative use of formulated product testing results in the aquatic organism risk assessment is limited in the EU to a consideration of exposure by drift introduction to a water body. This is consistent with the Agency's approach to only consider formulated product exposure under conditions of intentional application to aquatic environments and incidental application by pesticide drift. Risk assessments for terrestrial wildlife are usually conducted in the EU on the basis of active ingredient exposure models and active ingredient effects data. However, when pesticide formulations is a bait or a granular product, exposure and risk assessment will involve consumption of the entire product. In such cases, formulated product effects data may be used for the assessment. But such analysis would be exceedingly rare given that the formulated product effects testing requirement is seldom invoked for terrestrial wildlife.

EUROPEAN COMMISSION
HEALTH & CONSUMER PROTECTION DIRECTORATE-GENERAL

Directorate E - Food Safety: plant health, animal health and welfare, international questions
E1 - Plant health

Sanco/3268/2001 rev.4 (final)
17 October 2002

Working Document

Guidance Document on Aquatic Ecotoxicology

in the context of the Directive 91/414/EEC

Table of Contents

1. Introduction

1.1 Status and purpose of this document

This document on aquatic ecotoxicology was conceived as a working document of the Commission Services and was elaborated in co-operation with the Member States (MS). It is intended to provide guidance for notifiers in the context of the review of active substances under Council Directive 91/414/EEC. It is not intended to prejudice the authority of MS in national authorizations. Further, this document does not preclude the possibility that the European Court of Justice may give one or another provision direct effect in MS.

The purpose of this document is to provide guidance both to regulatory authorities and notifiers on the interpretation of the aquatic ecotoxicology sections of Annexes II, III and VI. Its aims are to promote consistency and transparency, and to enhance the efficiency of the review process.

Tools and techniques in ecotoxicological risk assessment progress rapidly. It is noted that it can be difficult for both notifiers or applicants - as well as regulators - to take such progress fully into account in their dossiers and assessment reports during on-going reviews. To provide a reliable framework for the review process and to avoid undue delays, the current version of this guidance document should therefore only be used for the review of existing active substances notified in the third phase of the review programme according to Regulation 451/2000[1] and subsequent phases. For new active substances, the document should be implemented with dossiers submitted from 1 August 2003. However, some flexibility may still be necessary during a transitional period. It will not always be possible to submit calculations and assessments according to the FOCUS surface water scenatrios within the timelines forseen. Decision making should also take into consideration that certain data requirements (e.g. full fish life cycle studies) which are now triggered, may not have been obvious to applicants or notifiers at the time of their notification or dossier submission. Likewise, if this appears justified in individual cases and facilitates decision making, the updated guidance may be considered also for substances submitted in earlier phases of the review programme.

Throughout the document, reference is made to reports from workshops or other scientific meetings. These are provided for information and should be used if appropriate. Also the Scientific Committee on Plants (SCP) has provided important guidance related to aquatic ecotoxicology in its opinions on individual substances and on a previous revision of this guidance document (SCP 1999[2]).

1 OJ L 55, 29.02.2000, p.25
2 http://europa.eu.int/comm/food/fs/sc/scp/outcome_ppp_en.html

1.2 Legislative background

Annexes II Section 8 and Annex III Section 10 of Directive 91/414/EEC set out the data requirements on ecotoxicolgy for the inclusion of an active substance onto Annex I of the Directive and for the authorisation of a plant protection product at MS-level. It should be noted that the introduction to both these sections provide useful information on the purpose and use of data submitted. Annex VI of the Directive includes the decision-making criteria for the authorisation of plant protection products at MS-level. Given that no other harmonised criteria are currently available, Annex VI should also be used in an appropriate way to decide whether Annex I listing of an active substance can be recommended. Over the last few years, issues related to aquatic ecotoxicology have been discussed at various meetings in the context of Directive 91/414/EEC. Several points of Annexes II, III and VI were identified during these discussions where expert judgement is required or where there is scope for different interpretation. This document attempts to address these issues.

It is clear that the data submitted must be sufficient to permit an assessment of the impact on non-target species. In order to fulfil this objective, tests additional to those outlined in Annex II and III may be needed in individual cases if there is a specific justification.

1.3 Protection aims

Any environmental risk assessment has two prerequisites:

- Definition of suitable assessment endpoints which are understood as formal expressions of the environmental values to be protected (SUTER 1993),
- Establishment of a certain level of protection which encompasses the acceptability of effects and the uncertainty linked to the prediction of effects.

The protection of species is a relevant assessment endpoint but difficult to evaluate and therefore not appropriate as a measurement endpoint. Due to the complexity of the matter, particularly when biodiversity issues are included, there are no agreed proposals on these points either in the scientific or in the regulatory community. In general, the sustainability of populations of non-target organisms should be ensured. Structural and functional endpoints should be regarded of equal importance.

Within the context of sustainability of our freshwater resources, the following unacceptable effects of contaminants are mentioned by Brock & Ratte (2001; see CLASSIC document) and should be considered when deciding about the acceptability of risk to non-target aquatic organisms. The reader is referred to the aforementioned paper for additional guidance:

Decrease in biodiversity
This concerns negative effects on:

- *Overall species richness and densities*

This may be expressed as the number of taxa, diversity indices (or scores of multivariate techniques) for the total community or for taxonomic or functional groups.

- *Population densities of ecological key species*

Ecological key species are species that play a major role in ecosystem performance, productivity, stability, resilience, e.g.,
- species that are critical determinants in trophic cascades (e.g. piscivorous fish; large cladocerans)
- species which are "ecological engineers" i.e., those that have a large influence on the physical properties of habitats (e.g. rooted submerged macrophytes)

- *Population densities of indicator species*

- species with a high "information" level for monitoring purposes
- species protected by law and regionally rare or endangered species

Impact on ecosystem functioning and functionality
This concerns negative effects on:

- *Water quality parameters (e.g. increase of toxic algae; oxygen depletion)*
- *Harvestable resources (e.g. fish)*

Decrease in perceived aesthetic value or appearance of the water body

- Disappearance of species with a popular appeal (e.g. dragonflies; waterlilles)
- Visual mortality of individuals of fish, frogs, water fowl and other vertebrates
- Symptoms of eutrophication (e.g. algal blooms)

In a limited number of cases, the use pattern of the compound includes direct application of the plant protection product into aquatic systems (e.g. in-crop areas like rice paddies or aquatic weed control uses). In these cases, unacceptable impacts on ecological function instead of biodiversity parameters should be the main consideration when effects on aquatic systems are assessed. For uses in rice, the relevant guidance document should be considered ("Guidance document on data requirements for active substances used in rice." SANCO/10/90/2000, in preparation).

1.4 Structure of this document

The document is divided into eight sections as follows:

1. Introduction
2. Data requirement: This section provides further information on the basic data requirements for an active substance and associated formulated product.
3. Exposure assessment: This section provides an outline of the exposure assessment (including consideration of the new FOCUS surface water exposure assessment methods; FOrum for the Co-ordination of pesticide fate models and their USe) that should be considered when carrying out a risk assessment.

4. Standard risk assessment: The preliminary risk characterization which permits identification of potential issue areas for further assessment.
5. Higher tier risk assessment: This contains possible approaches for higher-tier risk assessment.
6. Metabolites: This section provides an outline of the data requirements and exposure estimates required to enable an assessment for metabolites.
7. Risk management: This section provides information on a range of possible risk management options.
8. Other issues: This section contains general issues which are difficult to incorporate into other sections.

Where possible, the document provides examples which are aimed to help with interpretation of the recommendations.

2. Data requirements for active substances and formulations and their use in standard risk assessments

2.1 General issues in toxicity test design

2.1.1 Limit-tests

In principle, toxicity tests should be of a dose-response design. However, it is sometimes impractical to test at concentrations as high as those that are required for classification purposes. Furthermore, such high concentrations are often of limited relevance to the concentrations used in the risk assessment (which are generally considerably lower than the limits for classification). Whilst in some OECD guidelines a limit is given regarding a single maximum concentration to be tested, in others there are no such recommendations. Consequently further guidance on what is reasonable is needed in some cases.

For active substances and formulated products, concentrations up to 100 mg/l should be tested where no effects are determined at lower concentrations, and if no other recommendations are given in the annexes, the relevant OECD guidelines, or in this guidance document. This limit is consistent with other EU-regulations, and permits hazard classification and labelling of active substances and products.

If studies with metabolites are triggered (see Section 6), in principle the same limit as for active substances applies if the metabolite can easily be synthesised. For metabolites which are difficult to synthesise, a lower limit e.g. 10 mg/l would be acceptable. In special cases, acceptable limit concentrations will be considered to be those where the limit concentration tested is more than 1000 times the Predicted Environmental Concentration (PEC).

If older limit tests are available where the test concentration is lower than 100 mg/l, it may be necessary to repeat the study if the test result is directly relevant for the risk assessment (i.e. the test organism concerned is the most sensitive endpoint).

2.1.2 Poorly soluble substances

For poorly soluble substances, limit concentrations lower than 100 mg/l may also be acceptable (see "Draft-OECD Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures", OECD series on testing and assessment No 23, December 2000). Precipitation of the substance in the test medium should be avoided because data generated under these circumstances are usually highly variable. They may also over- or underestimate the toxicity of the compound when it is in solution or is reasonably well-dispersed in the test medium. It is generally not sufficient to test the maximum water solubility of the substance because this is usually determined in studies with pure water under sterile conditions. Attempts should be made to reach the maximum solubility level expected under the test conditions, using either an appropriate solubilizer, auxilary solvent or dispersing agent. For some compounds, the solubility in pure water is likely to be higher than in standard test media as this

already contains dissolved material. If, on the basis of these results a potential risk is identified (from the appropriate toxicity exposure ratio (TER)), further testing may be necessary. If the compound is very difficult to work with, there should be further discussions with the Rapporteur Member State (RMS) or another competent authority. Studies on the formulated product might also be an appropriate way to deal with poorly soluble compounds especially if no effects occur at the solubility limit. Another option is t to conduct tests in water-sediment systems (see section 5.4.2.1).

2.1.3 Analytical measurements

Annex II and III require biological testing to be supported by chemical analysis. The purpose of these measurements is to confirm that the organisms were exposed at the desired concentrations. In general, the recommendations of the relevant OECD-guidelines are reasonable and should be considered. In the case of flow-through tests, additional analytical measurements should be conducted between start and the end of a study. At least three test concentrations should be measured (usually the lowest, the middle and the highest) at the beginning and the end of the test. In long-term tests, measurements should be conducted at several timepoints between the study start and finish. In semi-static tests, the "new" and "old" test media must be analysed. For unstable compounds, substances which adsorb to glassware, or where the maximum water solubility lies in the range of test concentrations, it is usually necessary to conduct additional measurements to confirm the exposure concentrations.

For older studies that do not have appropriate analytical measurements , it may be necessary to repeat the study if the test result is directly relevant for the risk assessment (i.e. the test organism concerned is the most sensitive endpoint).

2.1.4 Calculation of test endpoints

Toxicity endpoints (LC/EC50, NOEC, etc.) should usually be calculated using nominal concentrations. This is because nominal concentrations are the most suitable for calculating TER values with maximum predicted environmental concentrations (PECmax). In studies where the initial measured concentrations are < 80 % or > 120% of the nominal, toxicity values should be presented as measured initial concentrations. These can then be used in the risk assessment in the same way as nominal concentrations. This approach is especially relevant for static tests.

If the measured concentrations are < 80 % or >120 % of nominal ones during the test, toxicity values should be presented additionally as measured concentrations and the mean measured concentration for the relevant test period should be used to express toxicity. If the measured concentrations are very low compared to nominal ones, the validity of the test might be questionable, and a justification for using such study should be required. In short summary, the following rules apply in general:

- if measured concentrations are > 80 %, then the nominal concentration can be used to express toxicity,
- if initial measured concentrations are < 80 %, then toxicity values should be expressed as initial measured concentrations,

- if measured concentrations in semi-static and flow-through systems fall graduately below 80 % during the test, then toxicity values should be expressed as mean measured concentrations.

2.1.5 Acceptable guidelines

Tests conducted in accordance with internationally-recognised guidelines (even if not specifically recommended in the Annex II or III) can be accepted if the guideline is comparable with those guidelines mentioned in Annex II or III. Tests with species mentioned in the aforementioned guidelines are in principle acceptable, although not all species are indigenous in Europe. It is the responsibility of the notifier to identify which annex points the data are intended to cover, to address any relevant deviation from the guideline specified in the annexes, and to justify why the data should be accepted.

In general, the notifier should submit all available data which may be relevant for decision making, although studies from the published literature often lack detailed description. If the notifier wishes to uses data from published literature or other sources to fulfil data requirements outlined in Annex II or III, then these can only be considered if a clear description of the method and a detailed presentation of the results are provided. Information on the specification of the test material used and Good Laboratory Practice (GLP) status of the facility which performed the study should be provided. The onus is on the notifier to justify why such data should be accepted.

2.2 Toxicity testing with fish

2.2.1 Acute toxicity tests (Annex II Point 8.2.2)

According to Directive 91/414/EEC acute toxicity data are always required for rainbow trout (*Oncorhynchus mykiss*) and a warm water fish species.

2.2.2 Long-term/chronic toxicity tests (Annex II Point 8.2.2)

Annex II point 8.2.2 state that a chronic (long-term) toxicity study must be carried out unless it can be justified that continued or repeated exposure is unlikely to occur. Long-term/chronic tests are important as they are the only measure of sub-lethal effects.

The definition of 'continued' exposure is important in order to decide whether long-term/chronic studies are required. A long-term/chronic test should be required if the DT50 from the water-sediment study for the concentration of parent compound in the water column is ≥ 2 days at an environmentally relevant pH in the range of 6 - 9. In practice, this means that chronic data are nearly always required. It should be noted, that short-term exposure may lead to sublethal effects which are not covered by acute toxicity testing. If there are such concerns, in special cases further evaluations might be needed (see section 5.4.2.1).

Multiple applications of a plant protection product could potentially lead to 'repeated' exposure. If the proposed use of an active substance involves more than one application per season, long-term/chronic toxicity data are required, unless the DT50 in the water phase is < 2 days and the notifier has clearly demonstrated that due to the length of the spraying interval, prolonged exposure will not occur. Where such conditions apply, the potental risk from repeated acute exposures should be addressed in the assessment report on a case by case basis.

Annex II Point 8.2.2 states that expert judgement is required to decide which test should be performed (test in accordance with OECD 204, 210, 215 or a fish full life-cycle (FLC) -test).

There are some reservations concerning the OECD guideline 204 (fish extended mortality test) because mortality is the only endpoint covered, and the exposure duration may only be 14 days. Furthermore, the developmental stage tested is not particularly sensitive. However, studies which have been conducted in accordance with this guideline in recent years usually include exposure for 21 days, with mortality, growth and behaviour as endpoints. In addition, the developmental stage of the rainbow trout to be tested is the same as recommended in the OECD guideline 215 for the 'Juvenile Growth Test', given that the weight of a 5 cm long rainbow trout (OECD 204) is in the range of 1 - 5 g (OECD 215). A combination of both guidelines is therefore considered most appropriate. Hence, the study should have a 28 day exposure duration and include survival, growth and behaviour as endpoints. In order to avoid unjustified animal testing, existing valid studies conducted in accordance with OECD 204 but lasting only 21 days can also be used to fulfil the data requirement.

For some active substances, the submission of data in accordance with Annex II Point 8.2.2.1 might not be sufficient to fully address the need for chronic toxicity data in order to complete the risk assessment. In these cases, the need for a fish early life stage toxicity (ELS) -test or a FLC-test should be considered. Guidance is provided in Annex II Point 8.2.2. on when such data should be required.

The trigger value of < 0.1 mg/l (acute LC50 for the active substance) stated for the ELS-test should also apply to the FLC-test. FLC-tests may be required where the BCF is >1000, the elimination during the 14 day depuration phase in the bio-concentration study is <95% or the substance is stable in water or sediment (DT90 >100 days). However, taking into account that this type of study is difficult to conduct and often the results do not differ significantly from the ELS-test, the FLC-test may not be required if only one or two out of the toxicity, bio-accumulation and persistence triggers are breached. If all three triggers are breached the test should be required. If effects on reproduction or the endocrine system could be anticipated (e.g. based on data from mammalian toxicology studies), the need for a FLC-test should should carefully considered (see section 8.3).

Annex II Point 8.2.2 also states that chronic toxicity studies are not required if 'a suitable microcosm or mesocosm study is available'. It should be noted though that

microcosm or mesocosm tests do not usually include the endpoint of chronic toxicity to fish. However, where valid fish data from a microcosm study (e.g. survival, growth, and behaviour) or mesocosm study (e.g. free living, reproduction data) are available then these may fulfil the requirement for data under Annex II Point 8.2.2.

2.2.3 Triggering of a fish bioconcentration study (Annex II Point 8.2.3)

A log P_{ow} > 3 should be used as a general trigger to require a fish bioconcentration study as stated in Annex II. Annex II also states that where it can be justified that exposure leading to bioconcentration is not likely to occur, a study is not necessary. Where bioconcentration is not expected because a substance is not stable in water, the study should not be required. This reflects the requirements of OECD 305 which is only considered suitable for 'stable' organic substances. Consequently, where the DT90 in the whole system is < 10 days (as determined in a water-sediment study), a fish bioconcentration study should not be necessary, unless the proposed use of the active substance includes multiple applications at intervals short enough to result in significant long-term exposure.

2.3 Studies with aquatic invertebrates including sediment-dwelling organisms

2.3.1 Studies with Daphnia (Annex II Point 8.2.4 and 8.2.5)

Under Directive 91/414/EEC, *Daphnia* is used as a representative invertebrate. Acute toxicity data are always required, and chronic data are also required if there is continued or repeated exposure to be expected. Chronic data are therefore required for compounds that are applied more than once per season, or for those whose dissipation rate (DT_{50}) in water is greater than or equal to 2 days (see Section 2.2.2). In practice, this means that chronic data are nearly always required. It should be noted that short-term exposure may lead to sublethal effects which are not covered by acute toxicity testing. If there are such concerns in special cases further evaluations might be needed (see Section 5.4.2.1).

In the preliminary risk assessment, uncertainty factors of 100 and 10 are applied to acute and chronic endpoints respectively to account for potential inter-species differences in invertebrate sensitivity and other sources of uncertainty. *Daphnia* is used as a representative invertebrate because of its ease of culture and testing, the availability of international acute and chronic guidelines (OECD 202 and 211), and its sensitivity to toxicants. A recent review paper (WOGRAM & LIESS, 2001) has clearly demonstrated that for organic chemicals including a range of pesticides, *Daphnia magna* is usually among the most sensitive species. Even when there are more sensitive groups, these are generally less than an order of magnitude more sensitive than *Daphnia*. The Annex VI trigger values for further assessment have also been validated in a major review study by BROCK *et al.* (2000 a and b) which compared sensitive endpoints from laboratory studies with insecticides and herbicides to the results of field studies. For these compounds, the Annex VI trigger values were clearly demonstrated to be protective for invertebrates when comparing with the NOEC and LOEC values found in micro- and mesocosm studies.

2.3.2 Studies with additional invertebrate species (Annex II Point 8.2.4 and 8.2.5)

For certain uses or compounds, studies on additional aquatic invertebrate species may be a core data requirement (as opposed to their use in higher-tier assessments – see Section 5).

Annex II 8.2.5 indicates that there is a requirement for studies on gastropod molluscs and insects if continued or repeated exposure is likely to occur. However, in general this requirement is limited to chronic tests whereas acute test (see Section 8.2.4) with gastropods and insects are only required if direct uses in waterbodies are intended.

An accepted international guideline for a chronic test on gastropods is not available currently. Furthermore, gastropod molluscs are generally significantly less sensitive than *Daphnia* (see WOGRAM & LIESS, 2001). Consequently, for uses where a direct application is made to water, the notifier should make a reasoned case as to why gastropod mollusc data should not be required. This could include acute toxicity data demonstrating the relative sensitivity of molluscs to the active substance. A chronic study should only be required if continued or repeated exposure is to be expected.

For herbicides and fungicides, *Daphnia* acute and chronic toxicity data (with their associated uncertainty factors) are suitably representative for aquatic insects and other invertebrates. For insecticides however, it should be carefully considered whether additional data on aquatic insects are required. Whilst for most insecticides, *Daphnia* have been demonstrated to be representative (BROCK *et al.*, 2000b), the toxicity of certain recent chemistries which have very specific, receptor-mediated modes of action (e.g. neo-nicotinoids) may not be well-represented by *Daphnia*. Information on the mode of action of insecticides (from efficacy and non-target arthropod data) should be considered before deciding whether testing on an insect species is required. If the toxicity of an insecticide to *Daphnia* is low (48 h EC_{50} > 1 mg/l, 21 d NOEC > 0.1 mg/l), this may indicate selectivity. An acute toxicity test should then be carried out with first instar (2-3 d old) *Chironomus riparius* (48 h water-only study). There is currently no guideline for such a study available and there is a need to generate it in due course, but in principle the tests should be conducted using similar methodologies as for *Daphnia*. The toxicity data from the most sensitive organism (*Daphnia* or *Chironomus sp*) should be used in the standard risk assessment for invertebrates and the usual triggers for further assessment applied. If a long-term/chronic study on insects is already available there is no need to require additionally an acute one.

If the 48h EC_{50} for *Chironomus sp* is at least ten times lower than the *Daphnia* 48 h EC_{50}, then a chronic study should also be conducted with *Chironomus sp* (see below). The provisions outlined in Section 2.1.1 apply for such tests. In these cases, the same triggers that are applied to *Daphnia* should be applied to the *Chironomus sp* data (*i.e.*,100 for acute toxicity, 10 for chronic toxicity to account for further potential differences in inter-species sensitivity of insects; see Section 4). These uncertainty factors may then be reduced by testing further species (see Section 5).

For insecticides which are insect growth regulators (*e.g.*, benzoyl ureas and similar classes), special consideration should also be given to the potential for effects on aquatic insects. Such compounds tend to have more pronounced effects over longer time periods than standard acute studies (due to their effect on moulting). Therefore, chronic studies with *Chironomus sp* should generally be conducted, unless it can be

clearly demonstrated that the onset of effects is rapid and that *Daphnia* are of similar sensitivity to chironomids.

2.3.3 Available data on estuarine/marine invertebrates (Annex II Point 8.2.4 and 8.2.5)

In some cases, data are available on estuarine/marine invertebrates (*e.g.*, *Mysidopsis bahia*, oyster embryo larval studies). At present, there is no requirement under Directive 91/414/EEC to perform these studies, but if data are available, they must be submitted and should be considered in the risk assessment. The notifier should make a reasoned case as to the relevance of data on estuarine/marine organisms to the risk assessment.

2.3.4 Tests with sediment-dwelling invertebrates (Annex II Point 8.2.7)

2.3.4.1 Introduction

Annex II point 8.2.7 states:
"Where environmental fate and behaviour data required in Annex II Section 7 report that an active substance is likely to partition to and persist in aquatic sediments, expert judgement should be used to decide whether an acute or chronic sediment toxicity test is required. Such expert judgement should take into account whether effects on sediment-dwelling invertebrates are likely by comparing the aquatic invertebrate toxicity EC50 data from Points 8.2.4 (acute) and 8.2.5 (chronic) with the predicted levels of the active substances in sediment from data in Annex III, Point 9 (Fate and behaviour in the environment)".

Additionally, Annex II Point 8.2.7 specifies *Chironomus* sp. (Insecta, Diptera, Chironomidae, Chironominae) as the required test organism to assess potential effects on sediment-dwelling organisms. Although the general triggering factors are identified, no trigger values are stated in Annex II and no testing guideline is specified. Further guidance on such triggering (taking into account partitioning, persistence, and potential for toxicity) and test methods is provided below.

2.3.4.2 Triggering of sediment toxicity tests with invertebrates

As indicated above, triggering for sediment studies should take into account the potential for exposure via the sediment, and potential for toxicity. For active substances, a test on sediment-dwelling organisms should be required if, in the water-sediment fate study (*e.g.*, OECD 308), >10% of applied radioactivity represented by the parent compound is present in the sediment at or after day 14, and triggers to identify potential risks to invertebrates for toxicity are met. For information on the triggering of sediment toxicity studies with metabolites or degradation products, see Section 6.6.3.

To prevent unnecessary testing with substances of low toxicity to invertebrates, the NOEC in the chronic *Daphnia* test (or in a comparable study with insects when this group of organisms is more sensitive) must be < 0.1 mg/l for testing on sediment-dwelling organisms to be warranted. This number was chosen because on the basis of data from monitoring studies it is unlikely that higher concentrations will often occur in surface waters. Furthermore, the use of toxicity values from *Daphnia* tests as trigger for requiring tests on sediment-dwellers is mentioned in Annex II. A recent review that compared toxicity data for *Daphnia* with that for sediment-dwellers supports the aforementioned approach (STRELOKE et al. 2002, see also section 2.3.2).

For persistent substances (see EU-Guidance-Document 9188/VI/97), it may be justified to require a life-cycle test on chironomids to generate data on effects on reproduction. However, a standardized test method is not available, and there have only been a limited number of studies published in the literature.

For compounds which do not reach the "10 % trigger" but are applied more than once during the season, due consideration should be given to the potential for accumulation of residues in the sediment. Exposure triggers based on the water-sediment study are more difficult to apply to such use patterns because in the water-sediment study, typically only a single application is made. However, the development of the FOCUS Step 2 calculator (Step1_2 in FOCUS – see section 3) now permits a TER-based approach to triggering to be applied.

At FOCUS Step 2, as well as including drift, potential inputs from the soil compartment (via drainage/runoff) are included. The compound is partitioned between 30 cm depth of water and 5 cm depth of sediment, and is degraded. At Step 2, it is assumed that both the soil and water compartments experience no dilution, and that an equilibrium develops between the sediment and water compartments, with concentrations only influenced by degradation.

It is well-established that for non-polar organic compounds of log Pow up to 5 that in such a system at equilibrium, adequate predicitions of toxicity in sediment can be made from the concentration in the water phase (DI TORO *et al.*, 1991). Because the FOCUS calculation partitions the compound between water and sediment and assumes that an equilibrium exists (worst-case because in nature dilution would be expected), the concentration in the water phase will reflect the 'bioavailable' concentration in the sediment. Consequently, using the appropriate water phase concentration, *Daphnia* toxicity data and the standard Annex VI triggers for invertebrates, it is possible to determine whether there is potential for sediment toxicity. Hence, if the TERs (based on the maximum exposure concentration at Step 2 from the 'Step1_2 in FOCUS' calculator) for *Daphnia* are less than 100 or 10 for acute or chronic endpoints, then testing of sediment dwelling organisms should be required, if the sediment exposure triggers are met. Some example calculations are included below (see Annex 1).

For insecticides where it is possible that *Daphnia* are not a representative test organism (see Section 2.3.2), acute toxicity data for *Chironomus riparius* can also be used to trigger long-term sediment studies. If the TER resulting from the maximum PEC at Step 2 and the *C. riparius* 48 h LC50 is less than 100, then long-term sediment testing is required, if the sediment exposure triggers are met.

2.3.4.3 Sediment-dwelling organism testing methods and endpoints

Although Annex II Point 8.2.7 specifies *Chironomus* sp. as the test organism, and survival and development (including emergence of adults) as endpoints, no further guidance is included on the type of study to be conducted. Two methods for testing sediment-dwelling organisms (in the presence of sediment) are available, both using *Chironomus* sp. The studies are quite distinct in that the first is a "spiked-sediment" toxicity test which expresses the results in terms of a concentration in the sediment (see draft OECD 218). The second is a "spiked water" toxicity test with sediment-dwelling organisms and expresses results in terms of a concentration in the water phase (see draft OECD 219).

There has been some debate about under which circumstances the "spiked water" or "spiked sediment" method is most appropriate. Data generated using either method should be judged on its own merits, although the spiked water test may be seen as providing a more realistic exposure scenario for most cases. However, data from spiked sediment studies can be particularly useful for addressing risks from exposure to contaminated sediment, particularly if there is an accumulation of the compound in the sediment over time (e.g., from multiple applications and/or via different exposure routes).

For sediment toxicity tests the concentrations in the pore water, the overlying water, and the sediment should be measured. There are some reservations with respect to the draft OECD 219 which includes the fact that analytical measurements in sediment are not routinely conducted. It can be argued that such analyses are not necessary if suitable data on the partitioning of the compound from a water-sediment study are available. In fact, studies with four radiolabelled substances showed that the partitioning in the water-sediment study and in the "spiked water" test with sediment-dwelling organisms should be comparable (STRELOKE&KOEPP, 1996). Therefore, reasoned cases which include estimation of likely levels in sediment, utilising data from the water-sediment study, may also be acceptable. In such situations, the notifier should demonstrate that the conditions in the water-sediment study are comparable to those in the "spiked water" test. The estimation of levels should include consideration of metabolites present in the sediment where this is relevant for the risk assessment. Additional analytical measurements in a study may sometimes be valuable to decide on the validity of a test and may help to avoid additional testing with living organisms.

NOEC values from "spiked water" studies that are expressed as initial concentrations in the water phase should be compared to initial PECs for the water column, and those from "spiked sediment" tests should be compared to PECs in sediment. Since both studies are long-term tests, the appropriate trigger for further evaluation is 10. If the trigger is not passed, a range of higher-tier studies are possible to further refine the risk assessment (see Section 5).

Toxicity to sediment-dwelling invertebrates may also be addressed in a suitably designed microcosm or mesocosm study.

2.4 Studies with Aquatic Plants (including algae and macrophytes)

2.4.1 Species for algae tests (Annex II Point 8.2.6)

A test with green algae is required in all cases. For herbicides, an additional test (conducted in accordance with internationally recognised guidelines) is required on a further algal species from a different taxonomic group. The second species should be from a group other than green algae, such as diatoms or the blue-green algae. Plant growth regulators should be treated in the same way as herbicides because they act on primary producers.

Comparisons between the endpoints growth rate and biomass have been made and came to the conclusion that biomass - or cell number – is usually the most sensitive endpoint (RATTE 1998; STAVELY 1999). Nevertheless both biomass and growth rate should be reported. As there is no clear evidence available to indicate which is the most relevant endpoint for the field situation the lower figure should be used in the risk assessment. Toxicity values should be based on the period of exponential growth.

2.4.2 Aquatic macrophytes (Annex II Point 8.2.8)

Annex II states that a test on higher aquatic plants (macrophytes) has to be performed for herbicides. Tests should be conducted with *Lemna* sp.. There is a suitable ASTM guideline, a draft OECD-guideline and an EPA guideline (draft OPPTS 850.4400) available, which should be used until the draft OECD guideline is finalized. Plant growth regulators should be treated in the same way as herbicides because they act on primary producers.

The number of fronds is the most important endpoint but if for example toxicity values for biomass or other endpoints are lower these may be used for the risk assessment if appropriate.

Where on the basis of the standard *Lemna* test a high risk to aquatic plants is identified (i.e. TER<10), the notifier should consider providing further information to demonstrate that the risk to higher aquatic plants is acceptable. It may be possible to obtain information on the mode of action, the importance of the different routes of exposure and the range of sensitivity from effects seen in terrestrial plant tests. Additional studies, using a range of aquatic plant species may be required for highly active compounds (see Section 5). Where the justification for an acceptable risk is based solely on a *Lemna* recovery study, the relevance to other aquatic plants which do not have the same capacity for rapid reproduction and/or for which the sediment route of exposure may be important, must be fully addressed.

If there is evidence from efficacy data or data on terrestrial plants that the data for *Lemna* are not representative for other aquatic plant species (e.g. auxin simulators which can be more toxic to submerged plants than for *Lemna*) additional data with other aquatic plant species may be required on a case-by-case basis. The test protocol for such studies should be discussed with the RMS or the competent authority because no internationally accepted guideline is available.

At present, laboratory toxicity methods with aquatic macrophytes taxa other than *Lemna* are at an early stage of development, and will require further research before it is possible to develop a harmonized guideline. A protocol using *Myriophyllum* is being developed. However, notifiers are advised to discuss the study design with the RMS.

2.5 Study requirements for formulations *(Annex III Point 10.2)*

2.5.1 Acute toxicity tests with the formulated product (Annex III Point 10.2.1)

Acute toxicity studies should not be required for every formulation. However, co-formulants and solvents in formulations may significantly increase or decrease the acute toxicity of the active substance and there is some difficulty in predicting which type of formulations are critical in terms of such interactions. If the formulated product contains more than one active substance, this also complicates the prediction of toxicity using data on the individual active substances to the extent that tests on such products are usually required. Acute toxicity data on a formulation also takes the toxicity of the co-formulants into account, as their toxicity will also be exerted in the tests.

If the active substance is more acutely toxic when it is formulated, TERs should be calculated on the basis of the data for the product (as stated in Annex III Introduction to Section 10 (vii)).

Annex III states that in principle, tests should be carried out on one species from each of the three groups of aquatic organisms (fish, aquatic invertebrates and algae). Where the available information on the active substance indicates that one group is clearly more sensitive, then tests on the most sensitive species of the relevant group should be carried out. In this context, the most sensitive group is defined as being at least 100 times more sensitive than the next most sensitive. If the least sensitive group is at least 100 times less sensitive than the most sensitive, then formulation data are not required on the least sensitive group. If the most sensitive species tested with the active substance is either *Lemna*, *Chironomus* or other species then these should be tested with the formulation. For poorly soluble chemicals, tests on the formulated product may be required for a group which does not show toxicity for the active substance at the solubility limit.

If the formulated product contains two or more active substances, and the most sensitive taxonomic groups for the individual active substances are not the same, formulation toxicity data are required on all three groups.

There is some scope for extrapolation of toxicity data between similar formulations. In addition, in some cases it may be possible to reliably predict the toxicity of a "simple" formulation from data on the active substance and information on the co-formulants. The notifier should justify such approaches in reasoned cases.

2.5.2 Microcosm and mesocosm tests (Annex III Point 10.2.3)

These data requirements are discussed in section 5. It should be noted that the data derived from microcosm and mesocosm tests although generated with formulated products are usually also most important for the evaluation of the active substance. Therefore in fact these data also pertain to Annex II.

2.5.3 Chronic toxicity tests with the formulated product (Annex III Point 10.2.4)

Annex III Point 10.2.4 states that laboratory chronic toxicity testing "may be required for particular plant protection products where it is not possible to extrapolate from data obtained in the corresponding studies on the active substance".

It is unclear, based on current knowledge, what criteria should be used to decide whether laboratory chronic toxicity data are necessary for a particular plant protection product. It can be argued that chronic formulation studies provide valuable information on sublethal effects from exposure to active substance andco-formulant interactions. However, further research and discussion is required on the fundamental question of whether formulations persist as formulations over longer time periods in freshwater ecosystems. At present, a refined chronic/prolonged exposure assessment cannot be carried out for a formulation as Annex III does not require a water-sediment study for formulated products. However, comparisons between the data from analytical measurements in the toxicity tests with the activesubstance and the formulated product may be helpful in deciding upon this question (see Section 3.3). Without a refined exposure estimate, a NOEC from a chronic formulation study can only be compared with the initial formulation PEC, which may lead to an overestimation of risk.

Even though some important areas are yet to be resolved, "day-to-day" decisions on the need for long-term/chronic data on specific formulations have to be made. The following guidance may be useful to address this issue case-by-case:

Long-term/chronic tests with the formulated product should be required for that group of organisms where the formulated active substance is more acutely toxic than the technical active substance by one order of magnitude or greater. Relevant information especially concerning the effect of specific coformulants on the fate and effects of the active substance could be required and used more routinely in the assessment. Further, the LC/EC50 in the acute test on fish or *Daphnia* with the formulated product must be <10 mg formulation/l. However, long-term/chronic tests with the formulated product are not necessary if continued or repeated exposure is not possible (see Section 2.2.2 and 2.3.1). Therefore, long-term/chronic toxicity data are not required if the notifier can clearly demonstrate that the formulation will not persist in natural water-sediment systems and that continued or repeated exposure will not occur.

In general, the same type of toxicity studies should be submitted as for an active substance. However, static tests may be more useful than flow-through studies as the former are slightly more relevant to the exposure which could occur in the field. An alternative is to conduct a specifically targeted microcosm study with the formulated product to address the long term risk.

There is scope for extrapolation of toxicity data between similar formulations. The Notifier has to justify such an approach in a reasoned case.

In the risk assessment, data from long-term/chronic tests with the formulated product should be used for TER calculations if these values indicate highest risk.

3. Exposure assessment

3.1 Exposure calculations and the implementation of FOCUS Surface Waters

The FOCUS Surface Water Scenarios Group was established in 1997 to redefine surface water exposure calculations for pesticide risk assessment in the EU. It was charged with developing a set of standardised modelling scenarios for drift and including drainage and runoff entry routes into surface water – a significant change to the *status quo*. The scenarios are based on a tiered sequence of exposure assessment *steps,* namely:

- Step 1 = Worst-case loadings.
- Step 2 = Worst-case loadings based on sequential application patterns (i.e. taking account of dissipation between applications).
- Step 3 = Realistic worst-case based on crop/climate scenarios (using realistic worst-case soils, topography, water bodies, climate, agronomy).
- Step 4 = Localised/regionalised risk assessment, including potential mitigation measures.

It should be noted that FOCUS is still under discussion and that the overview presented in this document might be amended subsequent to the finalization of the FOCUS report.

3.1.1 Step 1 and 2 Calculations

The scenario for the Step 1 and 2 calculations is a static ditch (no dilution from flowing water) of 30 cm water depth, and a 5 cm deep sediment layer is assumed with organic carbon content of 5% and bulk density 0.8 kg/l. A piece of software called "STEP1-2 in FOCUS" has been developed which allows the user to easily calculate Step 1 and 2 exposure values. Detailed documentation of the calculations is included with the software. A brief summary of the process is described below.

At **Step 1**, the application rate is assumed to be the maximum season's usage applied as a single dose, unless the DT50 in water for the compound is less than a third of the interval between treatments. In this case, the use rate for a single application should be assessed because there is no possibility of accumulation of residues in the ditch. Spray drift input is derived from the drift data of the BBA and is assumed to occur at the 90th percentile (benchmark value), varying with crop type. Inputs for aerial applications were derived from the US Spray Drift Task Force. It is assumed that the distance between the edge of the crop and the water body are fixed at 1 m for row crops and 3 m for tall crops. Run-off/erosion and/or drainflow are included at Step 1 as a single fixed loading of 15% of the application rate which occurs on the day of application. Outputs from the calculator include the maximum PECs in water and sediment, and then actual and time-weighted average PECs through time. The PECmax is the maximum predicted environmental concentration in water or sediment that is estimated to occur during the time course of application or thereafter (i.e. taking into account that for more persistent compounds there may be an accumulation of residues in water or sediment).

For **Step 2** calculations, a number of refinements are included to make the scenario more reasonable. Applications are assumed to be made sequentially at the rates and intervals specified on the use label. Degradation and partitioning then occurs between applications. Spray drift is considered separately for each treatment date, but the percentile for individual drift inputs is adjusted so that the overall probability of drift represents the 90th percentile loading (i.e. individual events for multiple applications are less than the 90th percentile). Distances between crop and water are the same as at Step 1.

At Step 2, interception of the soil deposit is also included, and this varies dependent on crop type and growth stage. Appropriate interception values are provided for a large range of crop types. Four days after the last treatment, a percentage of the residue remaining on the treated field (determined using the soil degradation rate) is then added to the ditch as a run-off/erosion or drainage input and is added directly to the sediment layer of the ditch. The magnitude of this loss is dependent on season (autumn, spring or summer) and region (North EU (N EU) or South EU (S EU)). Outputs from Step 2 are similar to those at Step 1. Time-weighted average and actual concentrations with time are calculated on the basis of a 'rolling window' approach, i.e., the maximum PECtwa across the whole exposure period is used (not just the PECtwa after the last application). Because under some circumstances, the PECs resulting from a single application may be higher than those from a multiple application (principally because drift inputs decline with multiple applications), the PECs resulting from a single application are also always calculated. The higher of the two should then be used in the risk assessment.

3.1.2 Use of Step 1 and 2 in the Risk Assessment Process

Appropriate PECsw and PECsed values generated by Step1-2 in FOCUS can be used to compare to toxicity values to generate TER values. If a compound fails either Step, the next level of exposure assessment is triggered. If areas of concern are identified at Step 2, there is no option to mitigate the exposure concentrations, for example by the use of buffer zones. The user is then required to perform the appropriate Step 3 calculations.

The rationale for this is that the assumptions made at Step 2 still represent very much a worst-case scenario. Step 1 and 2 are designed to identify compounds which are clearly safe, not to accurately quantify realistic risks under field uses which can then be adjusted according to mitigative practice. It is therefore inappropriate to apply mitigation at Step 2. Compounds that fail at this stage should be investigated with the more refined tools at Step 3 which can then be appropriately mitigated at Step 4 if concerns are not resolved. Furthermore, although the modelling process at Step 3 is significantly more complex than at Step 2, the calculations will be greatly facilitated by the software that has been developed to aid the user to both select the correct scenarios and run the appropriate models.

3.1.3 Step 3

Full details of the Step 3 scenarios and modelling approaches are included in the FOCUS surface water report. In very brief summary, Step 3 includes six drainage and four runoff scenarios. Each scenario represents soil and climate combinations for areas of the EU which are considered to be potentially vulnerable to drainage or runoff inputs to surface water. There are one or two of three possible water bodies (pond, ditch or stream) associated with each scenario according to local conditions, and each has a set of environmental properties and associated crops. The use pattern of the compound determines which scenarios are run (a "wizard" is available to assist the user with this). For each scenario, drift inputs are calculated with a spray drift calculator based on the BBA spray drift data, drainage inputs (where appropriate) are modelled with MACRO, and runoff inputs (where appropriate) with PRZM. The fate of the compound in the various water bodies is modelled with TOXSWA which has been modified for FOCUS to include dynamic hydrology. The outputs from each scenario modelled are similar to those at Step 2.

3.1.4 Step 4

In principle, Step 4 can be regarded as a higher-tier exposure assessment step. This may include a variety of refinement options of different degrees of complexity covering risk mitigation measures (no mitigation options are considered appropriate prior to Step 4), refinement of fate input parameters, or regional and landscape-level approaches. By its nature, Step 4 will be a 'case-by-case' process, depending on the properties of the compound, its use pattern, and the areas of potential concern identified in the lower tier assessments. As such, there are no specific recommendations for the Step 4 process. Rather, some general guidance on the sorts of approaches that may be applied is available. Additional scenarios than those proposed by FOCUS can be considered for risk mitigation purposes. The scientific validity of these scenarios must be supported by data and must be accepted on EU-level in a comparable way as the FOCUS scenarios.

3.2 Specific exposure scenarios

As noted above, specific scenario for rice is under development (SANCO/1090/2000). Also specific mediterranean crops such as olive trees, citrus or vineyards may require special scenarios, which must be considered case by case if these are within the major use relevant for Annex I listing.

For indoor uses, notifiers should provide a rationale as to whether these uses would lead to an exposure of aquatic organisms. The rationale should address the potential contamination of surface waters through drainage, condensation (inside of glass) and rainwater (outside of glass) from these facilities, and the potential risk to sewage treatment processes.

Currently, no harmonized approaches are available to determine the exposure of surface waters to plant protection products via volatilisation or dry deposition.

Recently, a FOCUS group on atmospheric transport of plant protection products (FOCUS-AIR) has been established and future guidance produced by this group should be considered in evaluations.

3.3 Use of time weighted concentrations (PECtwa)

The first stage of the acute and chronic risk assessments should be based on the initial/maximum PEC values. If the chronic TERs calculated using the initial/maximum PEC are below the relevant triggers, it may be appropriate to refine the risk assessment using PECtwa values if an unrealistic exposure regime prevailed in the relevant toxicity test.

In deciding whether the use of a PECtwa values is appropriate, fate and behaviour data, and the toxicity profile of the active substance (e.g. time to onset of effects in toxicity studies) must be taken into account. The notifier should present the time to onset of effects for the relevant endpoints. It should be recognised that the use of a PECtwa may overlook effects that result from exposure that occurred early on in the exposure period. In general, the use of PECtwa values in the acute risk assessment for fish and aquatic invertebrates is not appropriate because their use may lead to an underestimation of the risk resulting from the initial period of exposure. However, the use of PECtwa values may be relevant for the algae risk assessment since the primary endpoint in the algae toxicity study is growth rate inhibition over the whole exposure period (i.e. a sublethal parameter), rather than percentage of dead or damaged cells at the end provided nominal concentrations are maintained throughout the test. It should be noted that since the algae study is a static test (potentially including degradation during the exposure period) the use of a PECtwa should only be warranted if exposure under more natural conditions is predicted to differ significantly from that in the toxicity test.

If PECtwa values are used, particular consideration must be made of the potential exposure from metabolites, as these would not be taken into account in a PECtwa for the parent compound. In addition, the implications of multiple applications on a PECtwa should be considered. To assess risks for water-column organisms, PECtwa values should be derived from the degradation and dissipation in/from the water phase in the water-sediment study, rather than the DT50 for the whole system. The water-sediment study used in PECtwa derivation must be relevant to conditions in the field (eg in terms of pH).

PECtwa values should be compared with nominal concentrations from toxicity tests if measured concentrations show that test levels have been satisfactorily maintained over the exposure period (i.e. >80% of the nominal concentration) and in the water-sediment study the concentration in the water column fell below 80 % of the nominal concentration. It is recognised that for some active substances, it is very difficult to maintain nominal concentrations throughout the exposure period, even in an acute toxicity study (e.g. due to rapid hydrolysis). In these cases where reliable mean measured concentrations cannot be determined, it may be appropriate to compare the

LC/EC50 based on nominal concentrations with the initial PEC. If a reliable mean measured concentration of < 80 % of nominal can be determined, but there is clear evidence from the water-sediment study that exposure in the relevant toxicity tests is still unrealistic, then the mean measured concentration should be compared with an appropriate PECtwa.

For unstable active substances, where the toxicity data for the formulated product are relevant for the risk assessment, the exposure assessment is usually unrealistic because a water-sediment study with the formulated product is not available and therefore the PECtwa cannot be calculated. In such cases, it is possible to use the DT50 from the water-sediment study with the active substance, provided that the data from the analytical measurements in the toxicity tests with the active substance and the formulated product are comparable. If the formulated product contains more than one active substance, then it might be reasonable to use the same approach for the most toxic and/or persistent component of the product.

The use of PECtwa may not be appropriate for use with endocrine disrupting compounds since these effects may result from relatively short periods of exposure at critical developmental periods.

3.4 Ecological significance of exposure estimates

Point (iv) of the Introduction to Section 10 of Annex III states that "the final PEC estimations are to be adapted according to the different groups of organisms taking in particular into consideration the biology of the most sensitive species". Hence, the ECCO group "Ecotoxicology" should make sure that the final PECs are appropriate in terms of the biology and ecology of the most sensitive group of organisms identified when conducting the risk assessment for aquatic organisms. In addition, the exposure regime used in the relevant toxicity test and the time of onset of effects therein should be taken into account when deciding on the most relevant PEC.

4. Standard risk assessment

Annex VI C 2.5.2.2 states that
"Where there is a possibility of aquatic organisms being exposed, no authorisation should be granted if the toxicity/exposure ratio for fish and *Daphnia* is less than 100 for acute exposure and less than 10 for long-term exposure, or the algal growth inhibition/exposure ratio is less than 10, or the maximum bio-concentration factor (BCF) is greater than 1000 for plant protection products containing active substances which are readily biodegradable or greater than 100 for those which are not readily biodegradable, unless it is clearly established through an appropriate risk assessment that under field conditions no unacceptable impact on the viability of exposed species (predators) occurs - directly or indirectly - after use of the plant protection product according to the proposed conditions of use".

For groups of organisms not specifically mentioned in Annex VI, the appropriate TER trigger values for related groups should be used for acute and chronic risk assessments. For example, assessments using data on insects (including *Chironomus* sp.) should use the trigger values specified for *Daphnia* (acute or long-term, whichever is more appropriate). Currently, the TER trigger value specified for algae growth inhibition is also be applied to higher aquatic plants and bacteria.

Toxicity values from the "spiked water" study with sediment dwellers should be compared to surface water PEC values. Data from the "spiked sediment" study should be compared with whole-sediment PEC values. Care should be taken to use an appropriate PECsw for multiple applications. If the study design did not reflect the intended application pattern, use of a total load PECsw can be appropriate with a single spiked study.

5. Higher-tier risk assessment

5.1 Introduction

The scope of this section is to elucidate the "unless" clauses of Annex VI (see section 4) and hence provide guidance on the types of studies that can be undertaken to try to determine if "no unacceptable impact" occurs when the plant protection product is applied according to the conditions of use. At the time that Annex III was written, outdoor microcosm and mesocosm studies were the only higher-tier aquatic studies for which international guidance was available. Since that time, there have been substantial developments in the area of higher-tier effects assessments, and a range of approaches and studies are now recommended which can be used to refine the effects assessment. Consequently, the trigger values mentioned in Annex III should not automatically trigger a microcosm or mesocosm study, but should trigger a higher-tier effects assessment.

5.2 Higher-tier acute risk assessment

A number of uncertainties must be addressed to extrapolate from single-species laboratory data to a multi-species ecosystem. According to the Technical Guidance Document for Chemicals (EUROPEAN COMMISSION, 1996) the following has to be taken into account when choosing the appropriate uncertainty factor:

- Intra- and inter-laboratory variation of toxicity data
- Intra- and inter-species variation of toxicity data
- Short-term to long-term/chronic toxicity extrapolation
- Extrapolation of mono-species laboratory data to field impact on ecosystems

Whilst there is substantial data that demonstrates the uncertainty described by to the first three bullet points for plant protection active substances, there are only a few, if any cases, that support the uncertainty mentioned in bullet point 4 (i.e. the same species is more sensitive in a mesocosm study than in a laboratory test). The first two bullet points are also pertinent to the uncertainty factor of 10 prescribed in Annex VI. The uncertainty factor of 100 is therefore in general necessary to cover the above mentioned uncertainty resulting from the extrapolation from short-term to long-term/chronic endpoints. Since the overall level of uncertainty is lower if chronic data are available, an uncertainty factor of 10 should be used for the chronic risk assessment according to Annex VI of directive 91/414/EEC. However, it should be noted that the contribution of each of the different factors influencing the overall uncertainty can not easiliy be quantified and may differ in the field of acute and chronic testing.

In rare cases where the acuteto chronic ratio (A/C ratio) is low and the same PEC is used for acute and chronic risk assessment, the acute risk may appear to be higher than the chronic risk due to the greater uncertainty factor that is applied to the acute

assessment. From a scientific point of view, this is not logical. In such cases, the real difference between acute and chronic toxicity is lower than was anticipated when setting general uncertainty factors. Under these circumstances, the use of a lower uncertainty factor than 100 in the acute risk assessment should be considered.

5.3 Reduction of the relevant uncertainty factor if data from additional single species tests are available

The testing of more species reduces the uncertainty of the risk assessment attributable to inter-species differences in sensitivity (see also Section 5.6). It therefore permits a reduction of the uncertainty factor that is applied to the lower-tier data. If a considerable number of additional species was tested in valid studies, then it is possible that the uncertainty factors that are applied to the lowest toxicity value could be lowered by up to an order of magnitude. However, the full order of magnitude reduction is likely only to apply to acute risk assessments, *e.g.*, Annex VI TER trigger for acute risk to fish and aquatic invertebrates.

5.4 Design and conduct of higher-tier effects studies including microcosm and mesocosm studies (*Annex III Point 10.2.2*)

5.4.1 Introduction

5.4.1.1 General considerations

Annex III states that where the TER_{acute} is < 100 for fish and *Daphnia*, less than 10 for alga or $TER_{longterm}$ is < 10 for fish and *Daphnia*, expert judgement should be used to decide whether a microcosm or mesocosm study is necessary. Extensive international guidance on possible higher-tier approaches are described in the proceedings of the HARAP (Campbell *et al.*, 1999) and CLASSIC (GIDDINGS *et al.*, 2002) workshops. Higher-tier laboratory studies have also recently been reviewed by BOXALL *et al.* (2001). The reader is referred to these documents for detailed discussions, examples and literature references. The design of studies for higher-tier aquatic effects assessment should always be carefully considered on a case-by-case basis, and should take into account the findings of the standard risk assessment.

The term "microcosm" can be used for small-scale studies, whereas the term "mesocosm" generally refers to larger outdoor tests. Microcosm studies can be an effective compromise between standard laboratory tests and mesocosm studies. Mesocosm studies can examine effects of pesticides on communities of organisms under simulated field conditions. The general relationship between data from standard laboratory tests and micro- and mesocosm studies for herbicides and insecticides is reported by BROCK et al. (2000 a and b). For fungicides, this relationship still needs to be assessed.

5.4.1.2 Defining endpoints from mesocosm and microcosm studies

The data from microcosm and mesocosm studies should be used to determine a number of endpoints which can then be used further in the risk assessment (e.g. to derive an ecologically acceptable concentration (**EAC**) – see below). For the relevant taxonomic groups in the study, a no observed effect concentration at the community level ($NOEC_{community}$) should be derived using appropriate statistical techniques (e.g. Principal Response Curves). In addition, NOECs for populations of relevant organisms should be reported ($NOEC_{population}$). Where there are effects at the community or population level, the time taken for recovery to occur should also be reported.

The $NOEC_{community}$, the $NOEC_{population}$ and the time taken for recovery should then be used to determine a no observed ecologically adverse effect concentration (**NOEAEC**). The NOEAEC is defined as being the concentration at or below which no long-lasting adverse effects were observed in a particular higher-tier study (e.g. mesocosm). No long-lasting effects are defined as those effects on individuals that have no or only transient effects on populations and communities and are considered of minor ecological relevance (e.g., effects that are not shown to have long-term effects on population growth, taking into account the life-history characteristics of the organisms concerned). Different recovery rates may therefore be acceptable for different types of organisms. The NOEAEC can therefore be higher than the $NOEC_{community}$ or $NOEC_{population}$. Thus, if at a single test concentration effects were determined but recovery occurs and the effect is considered of no concern for the ecosystem sustainability, that concentration should be used as NOEAEC. Different NOEAECs may be derived from a study depending on the protection aim (e.g. in-crop versus off-crop area).

When a NOEAEC is derived for a particular study, all of the NOECs that are lower than the NOEAEC must also be presented in order to facilitate interpretation. The lack of ecological relevance of these NOECs must also be justified.

The NOEAEC may be used for a direct comparison with the relevant PEC if uncertainty has been reduced considerably and the result from the study is relevant for overall decision making. However, this will require clear knowledge of all relevant endpoints and long-term effects. Otherwise an appropriate uncertainty factor should be applied leading to the EAC which was defined at the HARAP workshop ("An ecologically acceptable concentration was defined by the workshop as being the concentration at or below which no ecologically adverse effects would be expected. Depending on the type of study, this can be defined either directly (*e.g.* from semi-realistic multi-species or field studies) or through the application of appropriate uncertainty factors (*e.g.*, with additional single-species tests). Expert judgement is needed in the derivation of an EAC."). While the NOEAEC is study specific, the EAC is derived from an overall evaluation of a compound. In concept it is comparable to the Predicted No Effect Concentration (PNEC) defined for other chemical types in the EU framework (eg industrial chemicals, biocides, veterinary medicines, feed additives). However, there is not too much experience with the use

of the PNEC in higher-tier risk assessments and clearer differences might emerge in future. Therefore both terms should be used in parallel for the time being.

5.4.2 Microcosm

5.4.2.1 Studies with realistic exposure conditions

The environmental fate properties of a pesticide can be an important factor in the mitigation of risk under realistic environmental conditions. If dissipation is rapid, risk assessments based on toxicity studies performed under constant exposure conditions may overestimate potential risks. As a complementary approach to the PECtwa described above (Section 3.3), it is possible to simulate such fate dynamics experimentally in higher-tier studies. Initial indications of the potential influence of exposure on toxicity may be derived for some chemicals (principally those that readily hydrolyse or substantially adsorb) by comparing the results of static and flow-through toxicity tests for the same endpoint. If apparent toxicity is significantly less in static tests, then fate processes may significantly mitigate risks under natural conditions. Modified exposure studies are appropriate to address both acute and chronic concerns.

One approach to modified exposure studies is to alter the test system to allow a certain environmental fate process to take place, e.g., by the addition of sediment to the test system to simulate adsorption or degradation, or by exposing the test system to natural light conditions to simulate photolysis. In "fate simulation" studies the method used should be justified on the basis of its relevance to realistic environmental conditions.

Currently, no test guidelines are available for testing algae, *Daphnia*, fish and aquatic plants in water-sediment systems. However, there is some experience with tests on sediment-dwelling organisms. In general, the test organisms should be inserted before the test substance is applied. The test material should usually be applied to the water column of the water-sediment system, but other types of exposure might be reasonable for special purposes. Deviating exposure regimes should be used with care because data may only be related to a single use situation (see discussion on the CLASSIC workshop). In such cases, the protocol should be discussed with the competent authority. It is often advisable to determine sublethal endpoints, even if only acute exposure is expected to reduce uncertainty for such critical substances. Even a short-term exposure may lead to sublethal effects and this kind of uncertainty is especially relevant in connection with these type of higher-tier studies. The influence and sensitivity of parameters such as the composition of the sediment, the sediment to water ratio, suitable organic carbon content of the sediment, sediment depth, optimal performance of standard species in such tests systems are not yet well-understood.. Further work is needed to develop specific testing conditions which, on the one hand, are representative of environmental conditions and, on the other hand, ensure that the potential risk is not underestimated.

5.4.2.2 Microcosm - indoor multi-species tests

Indoor semi-realistic microcosms tests are experiments in systems that intend to represent natural assemblages of organisms characterised by several trophic levels and that, at least for the larger part, are constructed directly with samples of natural ecosystems. Species covering a wide range of sensitivities and biological diversity can be included. In general, indoor semi-realistic microcosms can include micro-organisms, planktonic, periphytic and benthic algae, zooplankton, meiofauna, macro-invertebrates and, when large enough, also macrophytes.

Many of the fundamental issues relating to semi-realistic laboratory microcosms also apply to equivalent outdoor studies in mesocosms.

There are several advantages of indoor semi-realistic laboratory microcosm tests over outdoor field tests:

- They can usually be run throughout the year. However, since they are constructed in part with samples from natural ecosystems, they closely depend on seasonal availability of biological material.
- There may be potential for a higher level of control over the experimental conditions when compared with an equivalent field system.
- Compared with outdoor studies, set-up costs can be less for tests in laboratory microcosms. However, for a given study design, costs for biological and chemical analysis are similar to outdoor studies.

There are a number of potential disadvantages of semi-realistic laboratory microcosms over larger outdoor mesocosms which should be considered in the selection of an appropriate risk assessment tool:

- They do not usually allow realistic population densities of large organisms (e.g., fish, newts, frogs and nymphs of larger insects). What is more, if these animals are allowed to be present in indoor semi-realistic laboratory microcosms, they can unduly disturb the test system.
- Long-term effects and recovery of species with complex life-cycles may be difficult to determine in indoor test systems.
- There is a lower level of field realism compared to outdoor tests because natural fluctuations in climatic conditions usually are not covered (although these can be simulated).
- The number of micro-habitats present in indoor test systems is usually limited.
- Adequate sampling without overly disturbing certain populations (e.g., macro-invertebrates and macrophytes) can be problematic. Free living macro-invertebrates, however, may be sampled by means of artificial substrates, identified alive, and returned again in the test system. The biomass of rooted macrophytes and the abundance of sediment-dwelling macro-invertebrates usually can be assessed in an adequate way at the end of the experiment only. An alternative approach might be the use of *in situ* bioassays with representatives of these organisms that can be sampled more frequently.

5.4.3 Mesocosm - outdoor multispecies tests

5.4.3.1 Introduction

Mesocosms offer the same advantages as microcosms, but in addition, they usually include a wider range of species and generally offer a greater potential to assess the response at the population and, especially, the community level. Furthermore, natural fluctuations in climatic conditions enhance the level of field realism. In particular, they enhance the probability of recovery of some species through e.g. colonization. Clearly, the individual concerns arising from the use of a substance must be investigated, and the test design must be tailored accordingly (on a statistically sound basis). However, there is also an argument for some standardisation of a microcosm and mesocosm study design in order to make data for different substances more comparable and ease the interpretation of results.

Mesocosm studies are useful in risk assessments when laboratory studies (lower- and higher-tier) indicate potential risks and they should be designed to test specific hypotheses about ecological effects. Mesocosm studies should focus on population-level and community-level effects in order to derive an NOEAEC.

5.4.3.2 Guidance on test methods

An exposure-response experimental design with replication is clearly preferred to ease data interpretation. If possible, this should include the maximum PEC. The selected concentrations should generally be based on the expected effects and not only on the PEC.

Previously studies have generally attempted to simulate field exposure ("simulation" approach). Studies where the chemical is uniformly dosed into the water ("toxicological" approach) are preferred. They are often more easily interpreted and can be extrapolated to a variety of risk assessment scenarios.

Application of the test substance should be made in the period between spring and midsummer when the communities are in their "growth" phases. Within this timeframe, species richness and abundance are usually most suitable, and the potential time available to observe rates of recovery is long.

Due to the density dependence of numerous ecological phenomena, the evolution of small and large systems will be different. For example, the species richness is frequently positively correlated with the size of an experimental system. Due to the relation between functional and structural properties of communities and food webs and the size of the system, the response of a mesocosm to the contamination by a toxicant is not independent of its size. Self-sustainability of the test systems should been taken into account (CAQUET et al., 2000). In particular, the size/complexity of the experimental system should be sufficient to:

- Ensure the development and reproduction of the organisms which are being studied,
- Give sufficient refuges to prey to avoid elimination by predators,

- Make the recycling of nutrients possible,
- Ensure potential functional redundancy.

The possibility of recovery may also depend on the size of the systems since large systems may be more resilient than smaller ones to toxicant effects.

In general, when constructing a mesocosm, efforts should be made to introduce all the functional groups. This includes primary producers and the various levels of consumers, avoiding introduction of top predators that may greatly influence the system. Studying fish in mesocosms can present difficulties and needs to be carefully considered. When the invertebrate community is the principal endpoint of the study, it is recommended that free-living fish are not included.

Macrophytes are an important structural and functional component of shallow aquatic ecosystems, and in general should be included in micro- and mesocosm studies that aim to simulate these environments. If macrophyte communities are to be the principal endpoint of the study, special efforts are required to establish a diverse and representative community. Efforts should be made to emphasise on the use of macrophyte species with relatively low growth potential, otherwise an experimental system might be deeply altered in their response to contaminants (see CAQUET et al., 2000).

The notifier should indicate the precise location of the experimental units, and information should be given on the respective location of control and treated systems. The presence of neighbouring natural ecosystems in the immediate vicinity of the experimental area should also be roughly indicated, if it influences the potential for recolonisation of the mesocosm.

The level of identification should be as high as scientifically justified or practically feasible (recognizing that there are constraints on species identification, especially for smaller species). Special efforts should be made for those groups that are identified in lower-tier studies as potentially the most sensitive.

Univariate statistical methods are recommended for investigating effects at the population level, and multivariate methods are recommended for describing community-level effects. The Principal Response Curve (PRC) method is a suitable multivariate technique designed to analyse microcosm and mesocosm tests (VAN DEN BRINK& TERBRAAK, 1999). The statistical treatment of data is very important and the use of the aforementioned multivariate technique is recommended to gain insight into the often complex changes in community structure over time and the possible relationship with treatment. However, the outcome of such evaluations should be carefully checked in the light of the raw data especially for the most sensitive endpoints.

5.4.3.3 Evaluation of test results

When reviewing the results of mesocosm studies, all groups and species should generally be considered of equal importance, as it is difficult to identify the 'key' species. Structural and functional endpoints are in general of the same importance. Species structure is usually the principal endpoint. Functional endpoints alone are not considered appropriate for protecting biodiversity which is the most important assessment endpoint. Therefore, in general, differences in species composition at the end of the study between treated test units and untreated controls, represent an effect unless these differences can be explained in terms of natural or incidental variations in population and community development.

It is important that a sufficient number of populations were present in the study to reach valid conclusions with respect to the most relevant uncertainty factor. Usually there are a few species available with high abundances for which univariate statistical methods can be used. A second group of species occurs usually with lower abundances but mainly the data for the controls give a conclusive picture on the occurrence of these species in the study. Furthermore, a tendency of increasing effects with higher concentrations is detectable or clearly no effects in all treatments. These species are also important and the data can be evaluated with multivariate techniques. They are also relevant for a decision upon the uncertainty factor. However, there is a third group of species which are scattered about controls and treatments randomly with highly diverging abundances. These species are usually not relevant for the decision on the most appropriate uncertainty factor.

It is particularly important to consider those groups of organisms which were identified as the most sensitive in the standard risk assessment. For certain taxa or endpoints, effects observed in a field study may be considered acceptable, if with appropriate expert ecological judgement, it is considered that they would not pose significant ecological risks to natural aquatic ecosystems. In general, to demonstrate an acceptable level of effect from a particular treatment regime there must be evidence that the treated system and controls are in a comparable state at the end of the study. Test duration should be long enough to be able to observe recovery.

For a rough orientation – and to facilitate communication in workgroups - on the overall level of concern related to aquatic ecotoxicology, the following guidance for assessment of effects can be used, which was developed by BROCK et al., 2000 b (see also section 1.3):

Class 1: "effect could not be demonstrated"

- no (statistically significant) effects observed as result of the treatment, and
- observed differences between treatment and controls show no clear causal relationship.

Class 2: "slight effect"

- effects reported in terms of "slight" or "transient" and/or other similar descriptions, and
- short-term and/or quantitatively restricted response of sensitive endpoints, and

- effects only observed at individual samplings.

Class 3: "pronounced short-term effect"

- clear response of sensitive endpoints, but total recovery within 8 weeks after the last application, and
- effects reported as "temporary effects on several sensitive species", "temporary elimination of sensitive species", "temporary effects on less sensitive species/endpoints" and/or other similar descriptions, and
- effects observed at some subsequent sampling instances.

Class 4: "pronounced effect in short-term study"

- clear effects (such as strong reductions in densities of sensitive species) observed, but the study is too short to demonstrate complete recovery within 8 weeks after the (last) application.

Class 5: "pronounced long-term effect"

- clear response of sensitive endpoints and recovery time of sensitive endpoints is longer than 8 weeks after the last application, and
- effects reported as "long-term effects on many sensitive species/endpoints", "elimination of sensitive species", "effects on less sensitive species/endpoints" and/or other similar descriptions, and
- effects observed at various subsequent samplings.

The following suggestions about the translation of effect classes into NOECs and NOEAECs may be considered. If only effects related to class 1 were observed, the NOEC and NOEAEC are the same which is not the case for effects belonging to the other classes. With respect to class 2 effects, a NOEC and a NOEAEC should be determined although the values should often be the same. There is a need to explain that effects occurred, but that these effects were regarded for some reasons as ecologically not adverse. For effects in class 3, a clear difference between the NOECand NOEAEC should be determined. A NOEAEC cannot be determined if effects belonging to class 4 and 5 were observed. Whilst for class 4 effects, it may be possible to use other tools (see below), to show that effects are acceptable, this could be very difficult for effects belonging to class 5.

Intrinsic recovery potential mainly relies on resting stages present in the treated system itself (e.g. resting eggs of Cladocera or rotifers, algal spores). The importance of this phenomenon will frequently be dependent on the duration of the pre-exposure period since resting stages are naturally produced when climatic conditions become unfavourable. Therefore, if the systems experienced one or more autumn-winter-spring cycle before treatment, the abundance would be greater than for recently built mesocosm. The precise "history" of the systems should therefore be indicated by the notifier. Effects may be considered of low ecological significance if recovery takes place in a given time period like 8 weeks, but this period should not be used as strict trigger because recovery depends very much on the life history of the species. Even if recovery is observed in a mesocosm study, the extent and rate of recovery has to be

considered in the context of natural aquatic systems and the proximity of unaffected sites to those affected.

Where recovery of a species is not observed, or is only incomplete in a mesocosm study, it is the responsibility of the data submitter to discuss this observation and explain how this relates to the likelihood of recovery in natural aquatic ecosystems. Furthermore, some species cannot recover in mesocosm studies simply because of the conservative study design (e.g. gammarids). It is recommended that additional tools (e.g. further laboratory studies) are used to address the remaining uncertainty. The replacement of species is not acceptable in general. But in some cases, the replacement of one species by another with a similar role in the ecosystem may be considered acceptable (e.g. for some algal species) if functionality is maintained and no further structural effects occur (e.g. no indirect effects on zooplankton). The replacement species, however, should have a similar function. For example a replacement of green algae by blue-green algae or photosynthetic-facultative flagellates is unacceptable. In any case, functional characterisation of mesocosms should be performed for a significant period of time since functionality may sometimes be maintained for a short-term period but may decrease later. The notifier has to provide clear evidence that the ecological function and community structure in the field situation is unlikely to be significantly affected. It is recommended that for all species affected in a mesocosm study, the likelihood of recovery under field conditions is fully addressed when evaluating the study results. All factors that may influence population/community recovery should be considered, and should include dispersal ability, life-history, breeding season, number of breeding attempts per season, abundance in the environment, spatial records, as well as the natural variability in population sizes and distributions.

Population-level evaluation of genetic properties should also be considered. Genetic variability is a matter of concern since spatially limited populations which develop in mesocosms may exibit significant differences in various characteristics (e.g. consanguinity, founding effects) as compared to natural populations of the same species. If the same experimental systems are used from one study to another, the case of selection of less sensitive genotypes cannot be excluded. In this case the evaluation of effects may be biased (underestimation of effects). Increased homozygoty may also alter the pattern of response of some species to pesticides. It is therefore recommended to replace sediment after an experiment before a new mesocosm study is started in the same testing facility.

Results of field studies should be accompanied by clear explanations as to why a given observed effect should be considered ecologically significant or acceptable when they are presented to regulatory authorities and that, wherever possible, such studies should be reviewed by groups of experts to provide the least-biased advice, although it is accepted that this may be difficult under current registration procedures. Connections could also be established with experts working in the field of biological conservation.

5.5 Risk assessment on the basis of higher tier data

5.5.1 Single species tests

The reason for conducting indoor single species tests even in water/sediment systems is usually to obtain more realistic toxicity values and not a reduction of the uncertainty factor. However in special cases if a considerable number of species was tested reductions are possible (see Section 5.3). Single species tests are usually not designed to address the potential for recovery.

5.5.2 Semi-realistic microcosm and mesocosm

Based on the experience with indoor semi-realistic microcosms so far, the uncertainty factors applied to results (i.e., NOEAEC, see section 5.4.1.2) of such tests need to be assessed on a case-by-case basis, taking into account the uncertainty and acceptability of the test. As an intermediate test, indoor semi-realistic microcosms may serve to highlight issues which need to be addressed in a future outdoor mesocosm test. Due to the generally smaller species diversity in indoor microcosms, pesticide-stress may lead to more or less exaggerated indirect effects, since in these less complex systems not all feedback mechanisms will take place that may dampen pesticide-stress in the field. In addition, more pronounced responses of sensitive populations may occur in indoor microcosm tests due to a slower dissipation of the pesticide from the water phase (eg, because of less-pronounced photodegradation) and the lower potential for natural recolonisation of eliminated populations. Nevertheless, indoor semi-realistic microcosm tests may be used to define an overall ecosystem effect level. There is, however, a need to define an NOEAEC and the subsequent EAC using expert judgement, as is the case for field studies.

It may be appropriate to compare an NOEAEC directly with the PEC, provided all the uncertainty has been satisfactorily accounted for. Otherwise, some uncertainty factor has to be applied to define the EAC (see section 5.4.1.2). The degree of uncertainty that is applied to these studies should be reduced in comparison to the uncertainty applied to the standard risk assessment but needs to be evaluated on a case-by-case basis and will depend on what other data are available in the risk assessment (useful guidance has been provided by the SCP in its opinion on esfenvalerate - http://europa.eu.int/comm/food/fs/sc/scp/out63_ppp_en.pdf).

It is proposed that the use of ecological models for extrapolation is developed further in the future. NOEAECs from reliable static mesocosm studies should be regarded as generally representative or possibly conservative for surface waters in most agricultural landscapes. Databases describing the abiotic and biotic conditions of surface water should be developed to aid interpretation and extrapolation between different waters and regions. Landscape ecology should be considered when evaluating the uncertainty of mesocosm results because water bodies in agricultural landscapes are often not isolated and/or completely exposed. However, in general other stressors than the use of the evaluated plant protection product should also be taken into account but currently there exists no guidance how to conduct these type of considerations.

5.6 Probabilistic Risk Assessment

Probabilistic risk assessment (PRA) is an emerging approach to environmental risk assessment, although it has been applied for many years in other scientific disciplines. Recently, the EC funded a workshop on this subject (HART 2001) which reviewed the 'state-of-the-science' and made recommendations regarding implementation and research needs. The reader is referred to the presentations in the EUPRA proceedings and the cited literature for a comprehensive view of current status as well as a discussion of the advantages and disadvantages of PRA (see http://www.eupra.com). There has also been a major review of probabilistic approaches in the USA under the Environmental Protection Agency ECOFRAM (Ecological Committee On FIFRA Risk Assessment Methods) project (information may be obtained from the EPA web site at http://www.epa.gov/oppefed1/ecorisk/). Further developments in this area of risk assessment are anticipated in the future.

In aquatic risk assessment, PRA can be applied in a variety of ways, at various levels of sophistication and complexity, covering both the effects and exposure aspects of the risk assessment. A range of these options are discussed in the EUPRA proceedings (see particularly Appendix 2 p 18 of the EUPRA report). PRA will usually be a tool for higher-tier risk assessment, and consequently the appropriateness of the risk assessments for addressing potential concerns will need to be considered on a case-by-case basis.

The traditional TER-based approach uses point estimates for the input parameters (e.g. lowest available toxicity figure, highest exposure level) and involves a global factor (= critical TER) to cover the various sources of uncertainty. Such a procedure may lead to an over-estimation of risk if the assessment is based on an extreme combination of several input values. Unfortunately, a deterministic assessment does not quantify whether that is actually true in a specific case. This problem could be overcome by probabilistic approaches. Performing a PRA involves assigning probability density functions to the various components that affect risk, and then carrying out Monte Carlo simulations or other calculations in order to estimate the probability that a certain event takes place. At present PRA has the following shortcomings:

- For many input parameters reliable information on the distribution is lacking;
- There are no common standard methods for the statistical calculations;
- Which effect percentage should be used ?

Some further criticisms of probabilistic approaches have been made by FORBES and FORBES (1993), namely:

- The need to describe species to a theoretical distribution;
- The assumption that the distribution of responses of species tested individually represents the effect on an ecological community;
- The assumption that the organisms selected for testing are an unbiased sample (an assumption of the statistical distribution);

- The need to generate larger amounts of data.

The advantages of using probabilistic approaches are:

- More of the available data are used than in a simple quotient approach;
- Through determining the shape of the sensitivity distribution, uncertainty associated with the linear extrapolations associated with standard lower-tier assessments is removed;
- The generation of additional data is encouraged, because generally more data provide a better definition of the distribution and a less conservative risk assessment.

Perhaps the most straight-forward application of PRA is the use of the "species sensitivity distribution" (SSD). In this approach, toxicity data are fitted to a statistical model in order to describe the distribution of sensitivities that would be expected in the "universe" of species. A review of such approaches has recently been published (POSTHUMA in HART 2001). SSDs have also been used by certain member states (e.g. The Netherlands). Some recommedations for the use of such approaches are also included in the HARAP workshop proceedings (CAMPBELL *et al.*, 1999).

The number and type of additional species that should be tested depends on what is known about the mode of action or selectivity of the pesticide. In general, for compounds which do not appear to be selective to aquatic organisms (*i.e.*, all standard tests organisms respond at similar - within an order of magnitude - concentrations), it is suggested that eight species could be used as a minimum to describe the distribution of sensitivities of aquatic organisms. Lower numbers may be appropriate for groups of organisms like fish which show a lower variability like for example algae. However, in cases where it is known that a specific group of organisms is particularly sensitive, then the species selected for further testing should be chosen from the relevant group (see also section 5.3).

5.7 Higher-tier risk assessment for compounds which have a considerable potential to bioaccumulate

5.7.1 Introduction

When the maximum bioconcentration factor (BCF) is greater than 1000 for plant protection products containing active substances which are readily biodegradable or 100 for those which are not readily biodegradable, a higher tier risk assessment should be conducted in accordance with Annex VI Point 2.5.2.2. As bioaccumulation processes often are slow and substances could be persistent, a chronic risk assessment is appropriate. The following exposure routes should be considered:

1) Direct long-term effects in fish due to bioconcentration;
2) Secondary poisoning for birds and mammals;
3) Biomagnification in aquatic food chains.

5.7.2 Direct long-term effects in fish

Additional studies on the chronic toxicity to fish might be necessary. The trigger values for the need for an ELS-test or FLC-test for fish should be applied (see Section 2.2.2) in principle.

- An ELS-test should be applied when 100 < BCF (whole body) < 1000 and the EC_{50} of the active substance < 0.1 mg/L. Result: long-term NOEC.
- A FLC-test is required when the BCF (whole body) > 1000, and the elimination of radioactivity during the 14 day depuration phase in the bioconcentration study is < 95%, the EC50 from an acute toxicity study is < 0.1 mg/l and the substance is stable in water or sediment (DT_{90} > 100 days). Result: chronic NOEC.

A simple worst case assessment can be conducted according to the following steps:

1) Take the appropriate PECwater (PECi or TWA) from environmental fate section.
2) Compare with the relevant long-term NOEC.

If the trigger of 10 is not met, a refinement of the risk assessment is necessary. This means that microcosm or mesocosm studies, which implicitly take into account bioaccumulation, should be submitted.

5.7.3 Secondary poisoning for birds and mammals

This aspect is discussed in detail in the guidance document on higher tier risk assessment for birds and mammals (SANCO/4145/2000) where more detailed guidance is given. A simple worst case assessment can be conducted according to the following steps:

1. Take the highest PECwater (TWA, 3 weeks) from environmental fate section;,
2. Take the whole body BCF for fish;
3. Estimate residues in fish: PECfish = PECwater * BCF;
4. Convert the residue (PECfish) to daily dose by multiplying with 0.12 (for mammals)., 0.21 (for birds) and compare with relevant long-term NOEL for mammals and. birds (expressed ad mg/kg/bw/d).

If the trigger of 5 is not met, a refinement of the assessment is necessary.

5.7.4 Biomagnification in aquatic food chains

For aquatic food chains, the substances of concern are those which have a potential for biomagnification, i.e. where the whole-body-residue in an animal at steady state is higher than the residue in its food (bioaccumulation factor, BAF > 1). It should be noted that in the long-term/chronic tests with *Daphnia* and fish, biomagnification is partly covered because test organisms eat contaminated food. For substances with such properties, exposure may increase along the food chain.

For persistent and bioaccumulating substances, a higher-tier exposure assessment should be conducted. To decide upon the need for this higher-tier exposure assessment, similiar trigger values as for a FLC-test for fish should be applied in principle, namely:

- The BCF (whole body) > 1000 and the elimination of radioactivity during the 14 day depuration phase in the bioconcentration study is < 95% and the substance is stable in water or sediment (DT_{90} > 100 days).

If these triggers are met, detailed food chain modelling (e.g. according to CARBONELL et al., 2000) should be performed, or microcosm/mesocosm studies, which implicitly take into account biomagnification, should be submitted. However, it should be carefully considered whether the models used are appropriate for the special type of exposure relevant for plant protection products. If a modelling approach is selected, a food chain including at least three steps (algae, algae-feeding-invertebrates, and invertebrates-feeding-fish) should be considered. The accumulation potential in algae can be estimated as a BCF for unicellular algae. For the accumulation through the food in invertebrates and fish, toxicokinetic equations for oral uptake and depuration like the following should be used

$BCF = F\alpha/kd$

Where F is the daily food intake, α the assimilation factor and kd the depuration constant.

For an initial assessment, default values representing worst case conditions for F and α can be used. The kd for invertebrates can be extrapolated from the kd for fish if the metabolic route is known and is also represented in non-vertebrate animals. An additional uncertainty factor can be required in some cases to cover the differences in the metabolic activity between fish and invertebrates. If a potential risk is identified using worst-case default values, single-species oral exposure studies (e.g. MUÑOZ et al., 1996) should be conducted.

For very persistent and bioaccumulating substances, appropriate higher-tier studies and predictive models must be specifically designed to properly address the potential biomagnification and bioconcentration risk. Studies and/or models must cover the potential risk associated with continued or repeated exposures at different trophic levels. Where appropriate for the organism and mode of action, due consideration should be given to the possibility of accumulation in certain target organs, differences in metabolic capacity among taxonomic groups, and the application of toxicokinetics suitable for addressing long-term exposure conditions. For extremely bioaccumulating and persistent substances it should be considered whether modelling and microcosm/mesocosm testing is appropriate at all because even the best test methods currently available may not be sufficient to fully investigate problems which are linked to these properties of a substance. The biomagnification risk is considered a key part of the assessment, and all potential exposure routes should be considered (e.g. vegetation residues, soil and substrates from greenhouse uses).

6. Metabolites

6.1 Introduction

The active substance of a plant protection product may be transformed in the environment by either abiotic or biotic processes. Under Directive 91/414/EEC, the potential risks that these metabolites pose to aquatic organisms must be assessed in certain cases.

The use of a pragmatic approach has been broadly supported in previous reviews of this issue. In its opinion (see SCP/GUIDE/023 – Final or http://www.europa.eu.int/comm/food/fs/sc/scp/out47_en.pdf) on metabolites, the SCP stated:

"As to the 10% trigger, the SCP supports this as a pragmatic screening approach. However, it is recognised that metabolites occurring at lower levels may well be ecotoxicologically relevant. Hence, all available information and expert judgement should be used to assess if metabolites <10% give rise to particular concern. Such metabolites should then also be subjected to a risk assessment rather than a specific justification."

6.2 Definitions

To facilitate clear understanding the following generic definitions are used in this guidance document:

1. Metabolite: for the purpose of this document, the term is used for all breakdown products of an active substance of a plant protection product, which are formed in the environment after the application, be it by biotic or abiotic processes;
2. Major metabolite: all metabolites that are formed in amounts of ≥10% of the applied amount of active ingredient at any timepoint evaluated during the degradation studies in the appropriate compartment (i.e. soil, water and/or sediment) under consideration;
3. Minor metabolite: all metabolites that are formed in amounts of < 10% of the applied amount of substance of active ingredient at any time during the degradation studies under consideration;
4. Ecotoxicologically relevant metabolite: a metabolite which poses a higher or comparable risk to aquatic organisms as the active substance. Such a metabolite is relevant for the overall decision on annex I inclusion or for definition of risk mitigation measures;
5. Definition of ecotoxicologically significant residues (Annex VI, B.2.6.2): an active substance or – if appropriate – a metabolite for which an analytical method has to be established for monitoring purposes (see section 8.1).

6.3 Potential routes of entry

Metabolites can contaminate surface water via the following main routes:

(a) An active substance can enter surface water via spray drift, volatilisation/ deposition (see Section 8.5), runoff and/or drainflow and then degrade in the water or sediment phase. The route and rate of such degradation is estimated in the water-sediment study.
(b) An active substance may degrade in soil and produce a mobile metabolite which may then enter surface water via drainflow or runoff. The formation of such metabolites is measured in laboratory and field soil transformation studies.
(c) Metabolites formed in soil may also enter groundwater and these then could be present in surface water where groundwater becomes or contaminates surface water. The formation of such metabolites is measured for example in lysimeter studies or calculated with appropriate modelling software (cf. FOCUS Ground Water Report).

These routes of exposure are the most relevant for assessing risks to surface-water organisms. Considering that exposure via such routes is assessed at the point of application (eg edge-of-field with minimal dilution in a drainage ditch), these represent a worst-case for the potential risks of metabolites to aquatic organisms.

Information on the extent of formation of metabolites in water-sediment or soil transformation studies as well as in lysimeter studies can then be used along with data on the properties of the metabolite (e.g. its adsorption and persistence) to model potential exposure concentrations in surface water. These data will be provided by the fate assessment.

6.4 Data Requirements

6.4.1 Fate section

According to Annex II Section 7, fate studies are required for all major metabolites. However, risk assessment is not restricted to these major metabolites but should also include minor metabolites. Concerns related to minor metabolites may sometimes be triggered by factors such as:

- a high likelihood that the metabolite will leach
- a consistent increase in concentration or percentage of applied radioactivity towards the end of a lysimeter, a soil metabolism or water-sediment study.

However, it must be noted that there are practical constraints to metabolite identification. In general, metabolites have lower molecular weight than their precursors, making identification and quantification increasingly difficult. Minor metabolites can be difficult to identify because of the unfavourable ratio of radioactive material per peak, and the possibility that soil constituents may interfere with both the chromatographic behaviour and the method of metabolite detection or

identification. Attempts to increase the amount of material analyzed by HPLC in order to compensate for the lower amount of radioactivity can result in chromatographic peak broadening or shifting in chromatographic retention and loss of resolution. There is therefore a low mass of material available to be purified (which results in further losses). The purification steps required to generate material suitable for mass spectroscopy (the best available method for metabolite identification) and further analytical work also inevitably lead to significant losses of radioactive material. Another difficulty with minor metabolites is their often transient nature, which depends partly on the dynamics of the microbial populations during incubation of the soils and their metabolic potential during the course of the study. These problems should be considered when decisions upon the technical feasibility of the identification of a metabolite are to be made.

6.4.2 Triggering of Aquatic Risk Assessments with Metabolites

If the metabolite is CO_2 or an inorganic compound, not being a heavy metal; or, it is an organic compound of aliphatic structure, with a chain length of 4 or less, which consists only of C, H, N or O atoms and has no "structures" or functional groups which are known to be of ecotoxicological concern, then no further studies are required and the metabolite is not considered to be ecotoxicologically relevant and is of low risk to the environment.

The following section outlines how the metabolites identified in the water-sediment study should be further evaluated, depending on their chemical and fate properties. If a metabolite is formed in either the sediment or water phase, then it should be considered whether toxicity testing is required. For many minor metabolites in particular, environmental fate data may not be available. However, it may be possible to estimate these parameters using the available information such as the structure of the molecule, its retention time, or its similarity to other molecules.

Metabolites may also occur in surface water via other routes or processes, e.g. they may formed in the soil and then pass to surface water. These metabolites should be assessed as follows:

1. If as a result of a soil degradation study a metabolite is formed, an assessment should be made as to whether it is likely to enter surface water via drainflow or runoff. Data in the fate package should enable generation of a reliable PEC and hence the risk for this type of metabolite should be assessed as for the parent active substance. The PEC for such a metabolite, together with an outline of its fate properties, will be provided by the fate assessment. If the potential exposure of surface waters is negligible, then further evaluation is not required. The need for effects data will depend upon these characteristics as well as the toxicity of the active substance (see below).

2. If a metabolite is formed via hydrolysis, it is feasible that the toxicity of the metabolite may have been assessed as part of the standard toxicity studies (see Section 6.6). Data from a hydrolysis study should also be used to decide to

which extent degradation and toxicity depend on the pH-value of the test medium.

3. If a metabolite is formed via photolysis, it is proposed to adopt a case-by-case procedure. For the time being, a rapid degradation of the active substance by direct photolysis in water (DT50 in the range of a few days when calculated for environmentally relevant conditions in June in Central Europe) together with the formation of a metabolite in amounts clear above 10 % may indicate that the metabolite is likely to occur under field conditions and, therefore, further testing (e.g. by a light/dark water-sediment study) should be considered.

4. If a metabolite is likely to occur in groundwater (for example, if it is measured in a lysimeter study), exposure of aquatic life may occur where groundwater becomes surface water including indirect exposure via drainage systems. Therefore, if as a result of appropriate fate studies or modelling, a metabolite (including non-identified radioactivity from lysimeter studies) is considered likely to contaminate groundwater, then an appropriate surface water PEC should be estimated to assess the risk to aquatic life. As a very worst case starting point, it is proposed that the PECgw is used as a PECsw. This should be determined by applying the FOCUS groundwater scenarios or by using the maximum annual concentration found in a lysimeter study. If concern is raised, then a more appropriate PEC should be determined. A dilution factor (to account for the dilution when the drainage water enters the waterbody) may additionally be considered which reflects the typical environmental conditions for the crops defined in the GAPs. In general, it is difficult to propose a number for the dilution factor because the concentrations of substances in drainflow/runoff and the properties of the receiving waterbodies vary considerably. To cover even realistic worst-case conditions, a factor of 10 may be considered reasonable in the first instance.

6.5 Calculation of Metabolite PECsw

Concentrations of metabolites in surface water and sediment can be readily estimated using the 'Step1_2 in FOCUS' software which has a specific module for metabolite PEC calculations. Inputs required are the percentage formed in both the soil and water-sediment transformation studies (i.e. accounting for both potential entry via routes in one PEC calculation), plus the Koc, DT50 and the molecular weight of the metabolite. If these values are not available, it may be possible to estimate the Koc on the basis of the estimated log P of the compound (this can be readily done with a variety of physical chemistry software), or by evaluating retention times on analytical columns (e.g. draft OECD 117, 121 and 122). Alternatively, a conservative Koc of 10 l/kg can be assumed with respect to the water phase. For a sediment assessment a conservative Koc of 10 000 l/kg can be assumed. Similarly, if no specific soil degradation studies are available on the metabolite, its decline in the studies where it was formed may be used to estimate a DT50. Alternatively, a conservative DT50 of 300 d may be used.

6.6 Requirements for Aquatic Organism Testing with Metabolites

As a general principle, it should be understood that data requirements raised in this context do not always have to be addressed by experimental studies. Notifiers are invited to address the open questions by any other available information in support of a scientific and rational assessment. Valuable sources of information include, but are not limited to:

- Consideration of molecular structure of the metabolite (active part intact?);
- The occurrence of metabolites in existing tests with the active substance or major metabolites;
- General knowledge on the relationship between the toxicity of metabolites and their parent substances;
- Available knowledge on related compounds.

Tests with metabolites may not be required where they are generated relatively rapidly by hydrolysis, as their toxicity may be exerted in the tests on the parent compound. In toxicity studies with intensive lighting (e.g. algae and *Lemna* tests), it could be assumed that metabolites which are formed as a result of photolysis are present in an amount which is relevant for field conditions and additional toxicity testing with metabolites detected in the photolysis study might not be warranted. This is particularly the case when static studies have been used. These conclusions should be supported by analytical measurements.

If more than one metabolite is considered significant, it may be sufficient to conduct only tests with the most important metabolite (highest amount, most comparable in structure with a.s.). Alternatively, an appropriately designed microcosm or mesocosm study to address the risk from the parent compound and metabolites could be undertaken. Again, analysis should confirm that levels of the metabolite were present in the system where organisms could be considered to have been exposed.

The principles for assessing metabolites should in essence be the same as those for active substances. However, unnecessary toxicity testing of metabolites especially with vertebrates should be avoided (see below). For major and minor metabolites which require experimental studies, acute toxicity tests with *Daphnia* and a single fish species and an algal study should be conducted. Metabolites should in general also be tested with *Lemna*, *Chironomus* or other species if these taxa have been the most sensitive with the active substance. Initially, it is only necessary to test the most sensitive species from a particular group (eg only rainbow trout if more sensitive than warmwater fish). Testing on additional species may be necessary where the risk to a particular taxonomic group is considered to be of concern and is predicted to be greater than that from exposure to the parent compound. If it can be demonstrated that certain taxonomic groups are clearly less sensitive to the active substance (by a factor of 100) than other groups, testing can be limited to those which are the most sensitive ones. If testing reveals that the toxicity of the metabolite to one taxonomic

group is similar to or higher than the parent then testing may be required on all taxonomic groups.

Recently the use of quantitative structure-activity relationships (QSAR) to evaluate toxicity of metabolites has been suggested (SINCLAIR&BOXALL, 2002). –These approaches are particularly useful for metabolites that no longer contain the toxophore, and they should be used, if appropriate. Especially when requiring toxicity tests with a minor metabolite or metabolites occurring in soil or lysimeter studies, the aforementioned aspects should be considered very thoroughly. From numerous tests with metabolites it can be concluded that in most cases metabolites are less toxic than the a.s. and therefore pose a lower risk than the active substance (STRELOKE et al, 2002; SINCLAIR&BOXALL, 2002). Whilst there are some exceptions (e.g. if the active moiety is still present in the metabolite structure or if known structures of higher toxicity are formed or if the metabolite partitions to a larger extent into sediment than an predecessor in the degradation procedure), it is very unlikely for toxicity to increase by more than a factor of 10. Taking into account that the PEC_{SW} for a minor metabolite is very often by a factor of 10 lower than the PEC_{SW} for the parent, it becomes obvious, that even the few cases where metabolites might be slightly more toxic than their parent substances, are covered by the following approach which should be used on a case by case basis.

In the first instance a PEC_{SW} for the minor metabolite should be calculated and compared with the most relevant toxicity value from the risk assessment on the active substance divided by 10 to cover the very unlikely increase of toxicity of the metabolite compared with the active substance. This TER should then be compared with the relevant trigger of Annex VI. In general, only toxicity tests with a minor metabolite should be required if this trigger is failed,. The same approach should be used for major metabolites formed in soil and those determined in lysimeter studies which may contaminate surface waters via groundwater, drainage systems or run-off.

In order to decide whether chronic testing is necessary, the intended uses, the fate and behaviour of the metabolite, and the acute TER values for the metabolite should be taken into account. In general chronic/long term tests are only necessary if the persistence trigger for chronic tests is surpassed for the metabolite (see sections 2.2.2, 2.3.1). In terms of the choice of taxonomic group(s) to be studied, this should take account of any acute toxicity data on the metabolite, the acute and chronic toxicity data on the active substance, and data on fate and behaviour in aquatic systems. Only if the metabolite is more acutely toxic than the active substance should long-term/chronic tests be required. Where acute toxicity data are available on fish and *Daphnia* for a particular metabolite, chronic testing should only be required on the more sensitive group. If in individual cases there is clear evidence that a metabolite is likely to be more toxic in chronic/long-term tests than the parent, or it exhibits endocrine disrupting properties, then chronic/long-term tests should be required with this metabolite.

For unstable active substances (i.e. those that do not meet the persistence criteria detailed in Sections 2.2.2 and 2.3.1), it may be more appropriate to conduct chronic

studies on the stable metabolite instead of the parent compound. For unstable active substances, where chronic toxicity data for the parent compound are not available and an environmentally significant metabolite exceeds the persistence criteria specified in Sections 2.2.2 and 2.3.1, chronic toxicity data should be submitted for this metabolite regardless of its acute toxicity.

In principle, for major metabolites found in the sediment of a water-sediment study, the same triggers should be applied to metabolites as for the active substance. That is, in order to justify testing, metabolites should be present and persist in sediment and have the potential to be toxic to aquatic invertebrates. Therefore to require testing, a metabolite found in sediment should be present in the sediment at a level of more than 10% of the parent applied radioactivity at day 14 or later. Clearly the potential to exclude testing on the basis of toxicity will depend on the data that is available for the metabolite. The notifier should therefore make a case as to whether sediment testing is justified based on what is known about the toxicity profile of the metabolite. For example, if risk assessments with *Daphnia* indicate that the potential risks are low, then no further testing should be required.

6.7 Risk Assessment for Metabolites

In principle, the risk assessment process for metabolites will be similar to that for active substances, albeit recognising that risk assessment cases will not always require specific study data for certain metabolites. If preliminary risk assessments indicate potential concerns then, as for parent molecules, risk refinement is possible either by refining effect concentrations or by refinement of the exposure concentration (see Section 5).

If higher-tier studies have been conducted with the active substance, or a relevant formulation, these studies may have also assessed the risk from the metabolites. It is advised that if a higher-tier study, e.g. mesocosm study, is being carried out then appropriate analysis should be conducted so that an assessment of both the exposure and effects of any metabolites can be made.

6.8 Defining ecotoxicological relevance

If as a result of the above risk assessment, a metabolite is considered to pose a similar or even higher risk to the aquatic environment than the parent active substance, and therefore, risk mitigation measures are needed, this metabolite is considered as "ecotoxicologically relevant". Such a metabolite – but also the active substance - must be included in the residue definition. For an ecotoxicologically relevant metabolite – but also the active substance - a concentration where no unacceptable effects on aquatic organisms are to be expected needs to be defined (see Section 8.1).

7. Risk management

It should be noted that a new FOCUS group on landscape ecology and risk mitigation has been established recently. The outcome of the discussions in this group should be taken into account when setting risk mitigation measures.

A standard risk assessment or even a higher-tier risk assessment (as referred to above) may indicate that the risk to aquatic life may only be acceptable providing that risk management measures are used. When an active substance is under consideration for Annex I listing, the RMS should include reference to possible risk management measures that are required to identify a "safe use". Decisions on appropriate risk management options should, however, be made at the MS-level, when plant protection products are registered.

The most obvious risk management or mitigation measure is a "buffer zone". This is an unsprayed strip between the target spraying area and a water body. This measure is currently used in several MSs. As well as buffer zones, a variety of other risk mitigation measures are available which could be applied by Member States to manage risk. Establishment of wind breaks such as rows of trees may reduce spray drift contamination (as recommended especially in the NL). Improved application technique may also minimises spray drift (as recommended especially in DE, UK, NL). Other member states like the UK allow reductions in the standard buffer zone where the application rate used is below the maximum approved rate. There may also be differences in risk to different types of water bodies, and since the FOCUS surface water group will recommend the inclusion of flowing waters in future scenarios, it may be possible to differentiate risk management measures for different water body types (e.g. ditches, streams, rivers, semi-permanent waters)

It may be necessary to consider scenarios other than the preliminary worst-case standard one (i.e. a small 30 cm deep lentic system completely contaminated by the maximum application rate). It may also be appropriate to use more suitable spray drift data which may take in to account environmental factors, or the use of low spray drift technology. Notifiers and/or MSs may submit additional scenarios which are representative for local use conditions for evaluation on EU-level. However, the suitability of these scenarios must be supported by data.

Examples from DE and UK and other MS are available on how to implement additional scenarios in current schemes of setting risk mitigation measures (MAFF 2000; FORSTER&STRELOKE, 2001):

8. Other issues

8.1 Definition of ecotoxicologically significant residues – aquatic life (Annex VI, 2.6.2)

With respect to the inclusion into the definition of ecotoxicologically significant residues – aquatic life - the active substance and ecotoxicologically relevant metabolites are always relevant (see Section 6.8). Regarding metabolites the hazard should additionally be considered and any hazardous metabolite should be included in the definition of residues (see below). Additional studies can be submitted to remove the metabolite from the definition of residues if it can be shown that the environmental hazard is low.

A monitoring analytical method (see guidance documents SANCO/3029/99 and SANCO/825/00) may be required for hazardous metabolites where the relevant toxicity value is lower than 100 mg/l. In case the metabolites are not likely to persist or bioaccumulate, the method will only be asked for if the relevant toxicity value is below 1 mg/l.The concentration of 1 mg/l covers even extreme concentrations measured in monitoring programmes together with an uncertainty factor. At the same time it is achieved that for a metabolite classified as "highly toxic" analytical methods are available.

The concentration where in accordance with Annex VI no effects on aquatic organisms are to be expected (relevant toxicity value for the most sensitive organism together with uncertainty factor) should be included into the "Definition of ecotoxicologically significant residues - aquatic life" which should be included into the list of endpoints for aquatic organisms.

8.2 Animal experimentation

For reasons of animal welfare, every effort should be made to avoid duplicate tests on higher animal species.

8.3 Endocrine Effects

The area of endocrine disruption is currently under a great deal of debate at both national and international level, and significant research efforts are underway to establish the importance of such mechanisms of toxicity to aquatic organisms. Endocrine disruption should be viewed as one of the many existing mechanisms of toxicity of chemicals and thus can be assessed within the normal conceptual frame-work. Endocrine concerns cover a potentially wide range of mechanisms, not just potential effects on reproduction, and so the development of regulatory procedures in this area is complicated. There is currently a significant amount of discussion underway at an international level (e.g. via OECD, US-EPA). These efforts are developing tiered-testing and risk assessment approaches, and it is therefore

premature to make firm recommendations at present until broader consensus on appropriate approaches have been agreed.

For the time being, evidence from appropriate mammalian studies should be reviewed to determine whether active substances are demonstrating potential endocrine effects. For example, thyroid or gonadal tumors, or effects on sex differentiation and sex organ development could be indicators of an endocrine effect. In such cases, it may be appropriate to use data from a fish ELS-test to assess potential for developmental effects, or from a fish partial life-cycle study to assess reproductive effects (although guideline methods for the latter are not yet available) as a first step. The need of further testing (e.g. FLC-test, *Xenopus laevis* test) should be considered.

Whilst endocrine disruption is an emerging area of science, there are no indications that it is different in terms of uncertainty to any other mechanism of toxicity. As such, in most cases endocrine disruption mechanisms of toxicity should be able to be dealt with within the same framework as other expressions of effect. This opinion is reinforced by the conclusions of the Report of the Working Group on Endocrine Disrupters of the Scientific Committee on Toxicity, Ecotoxicity and the Environment (CSTEE of March 1999).

Formulants clearly identified as endocrine disrupter like for example nonylphenol should not be used in formulated products. For such type of substances, a full data set including chronic/long-term tests should be required.

8.4 Organisms dwelling in groundwater

Increasing research on the complex of the biological ground water community has led to the recognition of the ground water ecosystem as a subject of protection in its own right (HEALTH COUNCIL OF THE NETHERLANDS , 1996; FRAUNHOFER-INSTITUT FÜR UMWELTCHEMIE UND ÖKOTOXIKOLOGIE, 2001). Therefore active substances and metabolites occurring in ground water (i.e. metabolites detected in lysimeter studies and metabolites from soil degradation studies for which entry into ground water has been predicted by means of FOCUS calculations) may be assessed with regard to their impact on ground water ecosystems in future. In the absence of more specific information and agreed testing guidelines, it can be assumed that ground water organisms are of comparable sensitivity as taxonomically and physiologically related surface water organisms. Crustaceans represent the most important ground water taxa and – from a preliminary scientific point of view - data on surface water crustaceans are considered adequate and sufficient to cover ground water organisms. However, recovery observed in higher-tier tests may not be relevant for organisms dwelling in ground water. If active substances have very specific modes of action and if crustaceans are not well-represented by *Daphnia* further considerations might be needed (see for example Section 2.3.2). Currently there are no agreed schemes for exposure and risk assessments available. Therefore this issue

should be further researched and, if appropriate, incorporated into revisions of 91/414/EEC and of this guidance document.

9. References

BOXALL A, BROWN C, BARRETT K. (2001). Higher tier laboratory aquatic toxicity testing. Cranfield Centre for EcoChemistry research report No. JF 4317E for DETR, 70 pp.

BROCK T C M; LAHR J; VAN DEN BRINK P J (2000a). Ecological risks of pesticides in freshwater ecosystems. Part 1: Herbicides. Alterra-Rapport 088. 124 pp., Alterra Green World Research, Wageningen, Netherlands.

BROCK T C M, VAN WIJNGAARDEN R P A, VAN GEEST G J (2000b). Ecological risks of pesticides in freshwater ecosystems. Part 2: Insecticides. Alterra-Rapport 089. 142 pp., Alterra Green World Research, Wageningen, Netherlands.

CAMPBELL, P.J.; ARNOLD, D.J.S.; BROCK, T.C.M.; GRANDY, N.J.; HEGER, W.; HEIMBACH, F.; MAUND, S.J.; STRELOKE, M. (1999): Guidance Document on Higher-tier Aquatic Risk Assessment for Pesticides (HARAP), SETAC-Europe Publication, Brussels.

CARBONELL G, RAMOS C, PABLOS MV, ORTIZ JA, TARAZONA JV (2000). A system dynamic model for the assessment of different exposure routes in aquatic ecosystems. Sci Total Environ. 247:107-118.

DI TORO, D.M., C.S. ZARBA, D.J. HANSEN, J.J BERRY, R.C. SWART, C.E. COWAN, S.P. PAVLOU, H.E. ALLEN, N.A. ALLEN AND P.R. PAQUIN. (1991). Technical basis for examining sediment quality criteria for non-ionic organic chemicals using equilibrium partitioning. Environ. Toxicol. Chem. 10: 1541-1586. ECETOC (1996). The role of bioaccumulation in environmental risk assessment : The aquatic environment and releated food webs. Technical report no. 67, ISNN-0773-8072-67, Brussels.

EUROPEAN COMMISSION (1996). Technical Guidance Document in Support of Commission Directive 93/67/EEC on Risk Assessment for New Notified Substances and Commission Regulation (EC) No 1488/94 on Risk Assessment for Existing Substances Part IV (97EU1-1/9)", European Commission, 1996, ISBN 92-827-8014-7.

FORSTER, R.; STRELOKE, M. (2001).Workshop on Risk Assessment and Risk Mitigation Measures in the Context of the Authorization of Plant Protection Products (WORMM). Mitt.Biol.Bundesanst.Land- Forstwirtsch. Berlin-Dahlem, Heft 381.

FRAUNHOFER-INSTITUT FÜR UMWELTCHEMIE UND ÖKOTOXIKOLOGIE (2001). Ökotoxikologische Prüfung von Pflanzenschutzmitteln hinsichtlich ihres Potentials zur Grundwassergefährdung - Ecotoxicological testing of pesticides with respect to their potential of endangering groundwater communities, UBA-Text 76/01.

GANZELMEIER, H.; RAUTMANN, D.; SPANGENBERG, R.; STRELOKE, M.; HERRMANN, M.; WENZELBURGER, H.-J.; WALTER, H.-F. (1995). Studies on the spray drift of plant protection products. Mitt.Biol.Bundesanst.Land- Forstwirtsch. Berlin-Dahlem, Heft 305.

GIDDING, J.M.; BROCK, T.C.M.; HEGER, W.; HEIMBACH, F.; MAUND, S.J.; NORMAN, S.M.; RATTE, H.T.; SCHÄFERS, C.; STRELOKE, M. editors (2002).Community-Level Aquatic System Studies – Interpretation criteria. Pensacola, FL, USA: Society of Environmental Toxicology and Chemistry (SETAC). p 60.

HART, A. (2001). Probabilistic risk assessment for pesticides in Europe: Implementation & reasearch needs. A report from the European workshop on probabilistic risk assessment for the environmental impacts of plant protection products, The Netherlands, June 2001. Central Science Laboratory, Sand Hutton, York, UK (ISBN 185 945 010 5).

HEALTH COUNCIL OF THE NETHERLANDS: COMMITTEE ON PESTICIDES AND GROUNDWATER (1996). Risks of pesticides to groundwater ecosystems. Rijkswijk: Health Council of the Netherlands, 1996; publication no. 1996/11E. ISBN:90-5549-135-7.

MAFF (2000). Local Environmental Risk Assessment for Pesticides – A practical guide, MAFF, London.

MUÑOZ, M.J., RAMOS, C. AND TARAZONA, J.V. (1996). Bioaccumulation and toxicity of hexachlorobenzene in *Chlorella vulgaris* and *Daphnia magna*. Aquatic Toxicology 35:211-220.

OECD (1999). Guidance document on aquatic toxicity testing of difficult substances.

RATTE, T. (1998). Influence of the groth pattern on the EC50 of cell number, biomass integral and growth rate in the algae growth inhibition test. Report für the German UBA (Project 36003010).

RAUTMANN, D; STRELOKE, M., WINKLER, R. (2001). New basic drift values in the authorisation procedure for plant protection products. In Forster, R.; Streloke, M. Workshop on Risk Assessment and Risk Mitigation Measures in the Context of the Authorization of Plant Protection Products (WORMM). Mitt.Biol.Bundesanst.Land-Forstwirtsch. Berlin-Dahlem, Heft 381.

SCP (1999). Opinion of the Scientific Committee on Plants on the Draft Guidance Document on Aquatic Ecotoxicology. SCP/GUIDE/023 – Final, 27 September 1999. http://europa.eu.int/comm/food/fs/sc/scp/out47_en.html

SINCLAIR, C.J., BOXALL, A.B.A. (2002) Assessment of the environmental properties and effects of pesticide transformation products. Report to DEFRA, Project No. PN-0930.

STAVELY, J.P (1999). An evaluation and comparison of the relative merits of various parameter used to express the results (at 72 and 96 hours) of algal toxicity tests. Report prepared for the EPA.

STRELOKE, M.; KÖPP, H. (1995). Long-term toxicity test with *Chironomus riparius:* Development and validation of a new test system. Mitt.Biol.Bundesanst.Land-Forstwirtsch. Berlin-Dahlem, Heft 315.

STRELOKE, M.; JOERMANN G.; KULA, H.; SPANGENBERG, R. (2002). Analysis of toxicity data on aquatic organisms for regulatory purposes. Poster at the SETAC-Europe meeting,Vienna.

SUTER, G.W. (1993). Ecological risk assessment. Lewis Publishers, Boca Raton, 538 pp.

VAN DEN BRINK, P.J., TER BRAAK C.J.F. (1999): Principal Response Curves: Analysis of time-dependent multivariate responses of a biological community to stress. Environmental Toxicology and Chemistry 18:138-148.

WOGRAM J. & LIESS M. (2001). Rank ordering of macroinvertebrate species sensitivity to toxic compounds by comparison with that of *Daphnia magna.* Bull. Environ. Contam. Toxicol. 67: 360-367.

10. Annex

10.1 Annex 1 – Worked examples regards sediment-dwelling organisms

1. Compound X is an insecticide which can be used 4 times at 14 day intervals at an application rate of 250 g ai/ha in vegetable row (arable) crops. In the water-sediment studies, Compound X was found at maxima of 21 and 30% after 14 days in the two sediments tested. It has a Koc of 1000, water solubility of 1 mg/l, and DT50s in sediment water system of 35 d and in soil of 14 d. The 48 h EC50 to *Daphnia* is 0.75 µg/l and the 21 d NOEC is 0.020 µg/l.

Using Step1_2 in FOCUS, the peak concentration of Compound X for four applications in the water phase is 4.7 µg/l and the maximum 21 d time-weighted average PEC is 0.73 µg/l. These can be compared to the *Daphnia* effect concentrations:

	Acute effect concentration (µg/l)	Long-term effect concentration (µg/l)
	0.75	0.020
Relevant PECsw	4.7	0.73
TER	0.16	0.027
Sediment testing triggered?	Yes	Yes

Sediment testing is triggered because the compound has significant sediment exposure and has demonstrated potential risks to aquatic invertebrates.

2. Compound Y is a fungicide which can be used 8 times a season at 7 day intervals at an application rate of 1000 g ai/ha in vines. Compound Y has a Koc of 850, a water solubility of 15 mg/l, and was found at maximum of 8 % in the sediment in a water/sediment study, where it degraded rapidly with a half-life of 8 days. The soil half-life was also rapid at 2 days. The 48 h EC50 for *Daphnia* of Compound Y is 3 mg/l and the 21 d NOEC is 0.1 mg/l.

Using Step1_2 in FOCUS, the peak concentration of Compound Y for eight applications in the water phase is 13 ug/l and the maximum 21 d time-weighted average PEC is 6 ug/l. These can be compared to the *Daphnia* effect concentrations:

	Acute effect concentration (mg/l)	Long-term effect concentration (mg/l)
	3	0.1
Relevant PECsw	0.013	0.006
TER	230	17
Sediment testing triggered?	No	No

Sediment testing is not triggered because although there is potential exposure in the sediment, the compound is of low risk to aquatic invertebrates.

10.2 Annex 2 - Worked examples regards metabolites

The following examples are provided as illustrations of how the risk to aquatic life from metabolites that occur at less than 10% of applied active substance may be assessed. It should be noted that these calculations are included only as examples, and should not be taken as precedent.

10.2.1 Approach using SINCLAIR&BOXALL, 2002

Compound X degrades in soil to two metabolites, one which occurs at >10% and one less than 10%. The key fate and ecotoxicological endpoints are outlined in Table 9.2.1 and 9.2.2. As this is a new compound, it is not known whether the toxicophore is present in metabolite B. As regards determining a PEC, no information is available on DT50 etc for metabolite B, however this metabolite occurs at a maximum of 4% of the applied active substance and therefore this information will be used when estimating the environmental concentration. The parent compound is applied once at 100 g/ha to cereals in the autumn. It is assumed that there is no interception.

Table 10.2.1: Key fate endpoints for compound X and metabolites

	Soil DT50 First order	**Koc**	**Water DT50 First order**	**Molecular weight**	**% found in soil**	**PECsw µg/l**	**PECsed µg/kg**
Parent active substance	15 days	615 ml/g	30 days	260	-	1.26	38.67
Metabolite A	69 days	240 ml/g	n.a.	270	40	1.02	12.28
Metabolite B	Default 300 days	default 10 ml/g for PECsw default 3000 ml/g for PECsed	n.a.	210	4	0.3	1.53

(NB PECs calculated using draft version of FOCUS Step 1. Calculated as outlined in section 6.5. May need to be amended once FOCUS draft is finalised.)

Table 10.2.2: Key aquatic endpoints for compound X and metabolites

	Fish 96 hr LC50 mg/l	***Daphnia magna* 48 hr EC50 mg/l**	**Alga 72 hr EC50 mg/l**	***Lemna* 7-day EC50 mg/l**	**Chiro-nomid 28-day NOEC mg/kg**	**Fish ELS NOEC mg/l**	***Daphnia magna* 21 day N0EC mg/l**
Parent active substance	0.25	0.67	0.01	10	>100	0.18	0.36
Metabolite A	>100	>100	2.3	n.a.	n.a.	n.a.	n.a.
Metabolite B	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.

From Table 10.2.1, it can be seen that a full set of fate endpoints are available for the active substance. For the Metabolite A there are some endpoints, whilst for Metabolite B no data are available and hence the default values as proposed in Section 5 have been used. From Table 10.2.2 it can be seen that a full set of acute toxicity studies have been submitted and that alga is the most sensitive species. Chronic toxicity data have also been submitted. Data have not been submitted on the toxicity of Metabolite A to *Lemna*. Data are also not available on the chronic toxicity of Metabolite A to fish or *Daphnia magna*. On assessment of the acute data on Metabolite A, it can be seen that the compound has low toxicity to fish and *Daphnia magna*, however it is still moderately toxic to alga. On the basis of these data it has been concluded that no chronic data on Metabolite B are required as alga are of primary concern.

No data have been submitted on the toxicity of Metabolite B. According to the approach provided above, where there is a lack of information regarding the presence or absence of the toxicophore and a lack of fate endpoints (namely Kow and dissociation constant), then the metabolite should be assumed to be ten times more toxic than the parent (see Section 6.6).

In Table 10.2.3 the TERs for the active substance as well as metabolites A and B are presented. On the basis of the available data on Metabolite A, it can be concluded that this metabolite is not 'ecotoxicologically relevant'.

It can be seen that by assuming the toxicity of Metabolite B is ten times higher than the parent, then the risk to aquatic life is still considered to be acceptable, i.e. TERs are above the relevant Annex VI trigger values. It can be further concluded that Metabolite B is not 'ecotoxicologically relevant'.

Table 10.2.3 TERs for active substance and metabolites A and B

	Fish	***Daphnia magna***	**Alga**	***Lemna***	**Chiro-nomid, chronic**	**Fish, chronic**	***Daphnia magna*, chronic**
PEC – parent (µg/l)	1.26	1.26	1.26	1.26	38.67 (µg/kg)	1.26	1.26
TER – parent	198	532	7.9*	7936	2585	143	288
PEC – metabolite A (µg/l)	1.02	1.02	1.02	1.02	12.28 (µg/kg)	1.02	1.02
TERmeta A	>98039	>98039	2255	n.a.	n.a.	n.a.	n.a.
PEC – metabolite B (µg/l)	0.3	0.3	0.3	0.3	1.53 (µg/kg)	0.3	0.3
TERmeta B	83	223	3.3	3333	6535	60	120

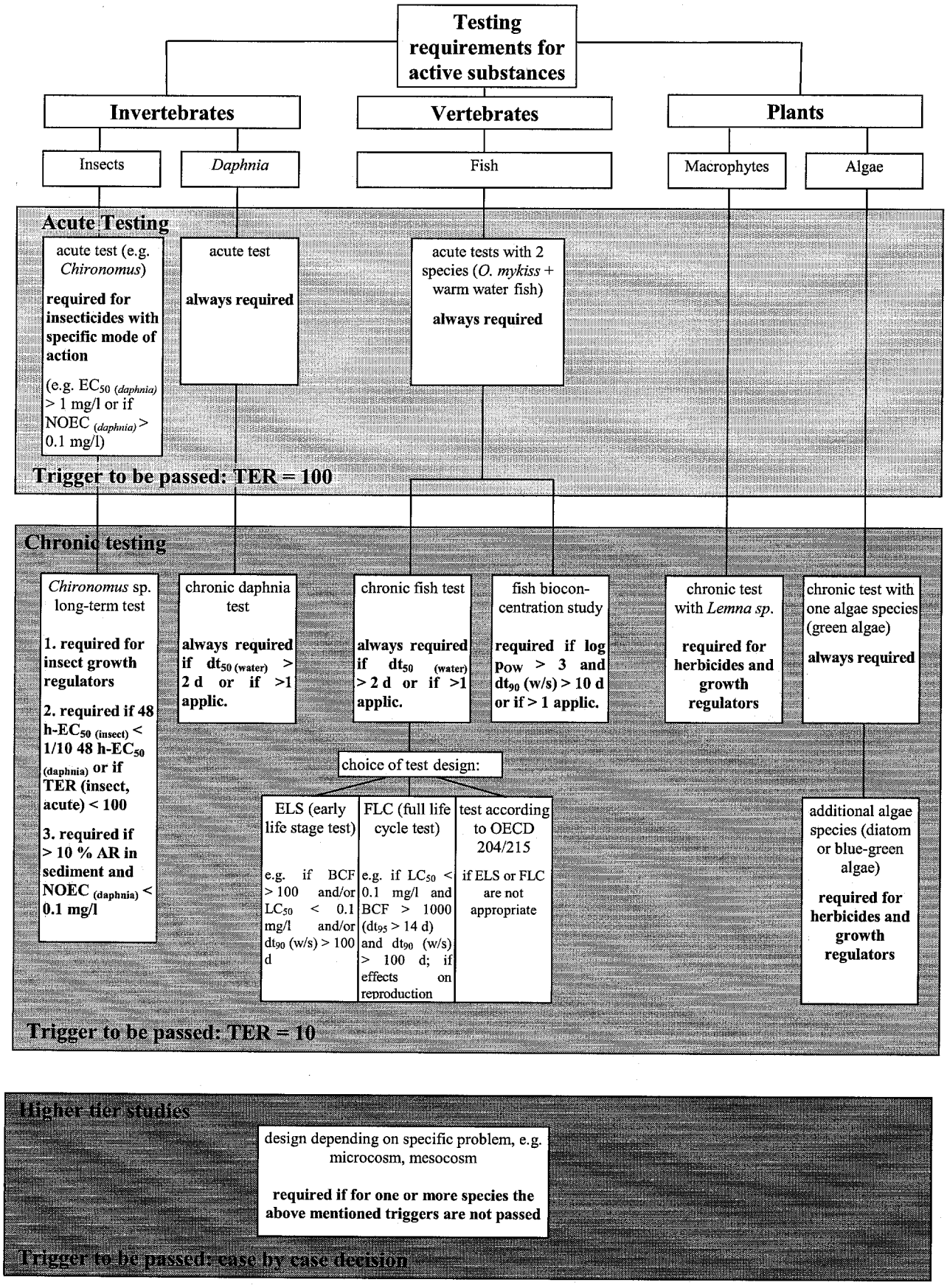
*** THIS TER IS BELOW THE APPROPRIATE ANNEX VI TRIGGER VALUE AND THEREFORE APPROPRIATE RISK MANAGEMENT MEASURES AND/OR HIGHER TIER DATA ARE REQUIRED ON THE ACTIVE SUBSTANCE.**

10.2.2 Example from the Scientific Committee on Plants (SCP)

The Scientific Committee on Plants (SCP) also has provided useful guidance. In the SCP opinion on imazasulfuron[3], the risk to aquatic life of ISPN was evaluated - a metabolite, which occurred at greater than 10%. However in assessing the risk from this metabolite a qualitative approach was used by the SCP. Full details can be found on http://www.europa.eu.int/comm/food/fs/sc/scp/out103_ppp_en.pdf. It should be noted that the SCP opinion was modified by the RMS in so far as the methods for calculating the PECs for the exposure route drainage were improved.

3 Opinion on the evaluation of imazosulfuron [th-913] in the context of Council Directive 91/414/EEC concerning the placing of plant protection products on the market. SCP/IMAZO/002-Final adopted 25 April 2001.

10.3 Annex 3: Testing requirements for active substances



Please note, that this assessment scheme is simplified. Additional tests are usually required for formulated products and metabolites. For more detailed information see the respective chapters in the text

EUROPEAN COMMISSION
HEALTH & CONSUMER PROTECTION DIRECTORATE-GENERAL

Directorate E - Food Safety: plant health, animal health and welfare, international questions
E1 - Plant health

SANCO/10329/2002 rev 2 final

17 October 2002

DRAFT
Working Document

Guidance Document on Terrestrial Ecotoxicology Under Council Directive 91/414/EEC

This document has been conceived as a working document of the Commission Services which was elaborated in co-operation with the Member States. It does not intend to produce legally binding effects and by its nature does not prejudice any measure taken by a Member State within the implementation prerogatives under Annex II, III and VI of Commission Directive 91/414/EEC, nor any case law developed with regard to this provision. This document also does not preclude the possibility that the European Court of Justice may give one or another provision direct effect in Member States.

1 Introduction

Article 5 of the Directive provides that "*in the light of current scientific and technical knowledge, an active substance shall be included in Annex I for an initial period not exceeding 10 years, if it may be expected that plant protection products containing the active substance ... do not have any unacceptable influence on the environment ...*".

Annexes II and III of Directive 91/414/EEC set out the data requirements for the inclusion of an active substance into Annex I of the Directive and for the authorisation of a plant protection product at Member State level. Annex VI of the Directive includes the decision making criteria for the authorisation of plant protection products at Member State level.

It is the purpose of this document to provide guidance to Rapporteurs, peer reviewing Member States, Notifiers and Applicants on the use and interpretation of the terrestrial ecotoxicology sections of Annexes II and III and to lay down agreed procedures and criteria for decision making. The general aim is to promote consistency and transparency in decision making and to describe agreed risk assessment procedures for the assessment of plant protection products in the context of the inclusion of their active substances in Annex I to Directive 91/414/EEC.

It has to be recognised that the authorisation of plant protection products after Annex I inclusion of active substances remains the responsibility of Member States. Risk management and risk mitigation measures described in this document do not pre-empt this authority of the Member States and are meant as a non-exhaustive list of agreed options, which can be taken into consideration on the Community level for decision making concerning Annex I inclusion.

The ecotoxicology data requirements for active substances and plant protection products are set out in Annex II, section 8 and Annex III section 10 of Directive 91/414/EC, respectively. It should be noted that the introduction to these sections provides useful information on the purpose and use of data submitted. It is clearly stated that the data submitted must be sufficient to permit a scientifically valid assessment of the impact on non-target species. In order to fulfil this objective, tests additional to those outlined in Annex II and III may be needed in individual cases if there is a specific justification.

Tools and techniques in ecotoxicological risk assessment progress rapidly and it is noted that it is difficult for both notifiers or applicants as well as reviewers to take such progress fully into account in their dossiers and assessment reports during ongoing reviews. To provide a reliable framework for the review process and to avoid undue delays, the current version of this Guidance document should therefore only be used for the review of existing active substances notified in the third phase of the review programme according to Regulation 451/2000[1] and subsequent phases. For new active substances the document should be implemented for dossiers submitted from 1 August 2003. However, some flexibility may still be necessary during a transitional period of 2-3 years. Decision making should take into consideration that certain higher tier data requirements (e.g. litter bag studies) which are triggered now, may not have been obvious to applicants or notifiers at the time of their notification or dossier submission. Likewise, if this appears justified in individual cases and

[1] OJ L 55, 29.02.2000, p.25

facilitates decision making, the updated guidance may be considered also for substances in earlier phases of the review programme.

The document is to be revised regularly, in order to reflect changes of test guidelines and of scientific knowledge.

2 General issues

2.1 Introduction to the assessment of chemicals in the terrestrial environment

The assessment of the effects and risks of chemicals for the terrestrial environment is a complex matter. This complexity comes, among others, from factors such as the need for sharing of the available landscape among urban/industrial activities, agricultural production in the form of agro-systems, and supporting terrestrial ecosystems. In addition, terrestrial systems are not associated with a single compartment, but with the interface between soil and the atmosphere. Although purely soil-dwelling organisms play a clear role, basic ecosystem functioning and biodiversity is associated with organisms, such as terrestrial plants, many invertebrates, and certain terrestrial vertebrates that are simultaneously or sequentially located in the soil or above-soil compartments.

The risk assessment for terrestrial ecosystems has been reviewed by the Scientific Committee on Toxicology, Ecotoxicology and the Environment (CSTEE 2000). According to this document:

"General adverse effects on the terrestrial environment include:

- Effects on soil functions, and particularly on the capacity of soil to act as substrate for plants including effects on seed germination, and those on organisms (invertebrates, micro-organisms) important for proper soil function and nutrient cycle conservation.
- Effects on plant biomass production, related to contamination of soil or air including deposition on plant surfaces. Plants are the source of food for the whole system (including humans) and have additional roles in terms of land protection, nutrient cycles, equilibrium of gases in the atmosphere, etc.
- Effects on soil, above-ground and foliar invertebrates, which represent food for other organisms, and cover essential roles as pollinators, detrivores, saprophages, pest controller, etc.
- Effects on terrestrial vertebrates exposed to contaminated food, soil, air, water or surfaces, with obvious economic and/or social consequences. Poisoned birds and mammals probably constitute the highest social concern, while reproductive effects, although less evident, represent a higher ecological hazard.
- Accumulation of toxic compounds in food items and through the food chain. Is a typical exposure route for animals within the contaminated ecosystem and represents an additional concern related to the consumption of this food by humans and domestic animals.

These concerns combine human and ecological interests. Direct human interests include managed species (cultivated plants and trees, bees, domestic animals) but also wild species essential as source for supplies (e.g. forest, pasture), landscape conservation (e.g. vegetation cover), or even for leisure (from gaming to bird-watching). From an ecological point of view, any of these effects will provoke a dramatic alteration of the structure and

functioning of the ecosystem which are considered the basic protection goals in ecological risk assessment."

However, as frequently noted by the Scientific Committee on Plants, the environmental risk assessment of plant protection products requires some adjustment of the generic ecological risk assessment framework as effects on living organisms considered as pests can be both acceptable and desirable.

Directive 91/414/EC includes the need for specific assessments on certain terrestrial non-target groups, such as terrestrial vertebrates, bees, other non-target arthropods, earthworms or soil micro-organisms, as well as additional generic assessments such as on soil macro- and mesofauna when triggered by fate properties (persistence).

Targeted risk assessment, using a combination of key ecological receptors and relevant exposure routes has been recently suggested as an efficient way of solving the complexity of the terrestrial environment risk assessment (Tarazona et al. 2002). This possibility fits perfectly with a protection aim established for plant protection products, allowing the identification of target species and non-target ecological receptors.

There is a common understanding that the ecological risk assessment aims not at individuals but at the protection of populations. In general the continuance of populations of non-target organisms should be ensured. Structural and functional endpoints should be regarded of equal importance.

2.2 Animal experimentation

For reasons of animal welfare all efforts should be made to avoid unnecessary tests especially on vertebrate species.

2.3 NOEC-values as summary parameters

In several tests the aim is to determine the no-observed-effect concentration (NOEC), a concept that has been challenged on scientific grounds (Laskowski 1995, OECD 1998). The OECD, and also ISO now give preference to regression-based parameters and in newly drafted guidelines give the choice for an ECx approach. (Note: The terminology referring to concentration (NOEC, and ECx) is used for convenience; the same applies, of course, to effect levels expressed as dose, application rate, etc.). NOEC tests are still acceptable, of course, however it should be ensured that the statistical power of the individual test is satisfactory. To that end some guidelines state the maximum permissible variation coefficient for certain variables. If such validity criteria are missing the typical power of that type of test should be used as a rule. For instance, if a test usually is able to detect a 20-% difference from the control then a treatment group with a difference of 40 %, which is statistically not significant, should not be accepted as a NOEC. For background information see OECD (1998). The OECD is currently working on a guidance document on statistical analysis of ecotoxicity tests.

2.4 Test substance, formulation testing

Test substance for Annex-II data requirements

In general the studies outlined in Annex II should be conducted using the technical grade material of the active substance. However, certain study types may be conducted with a

formulated product instead of the active substance. This may be applicable to, for example, non-target arthropod studies, the earthworm reproduction test and the soil micro-flora test. The formulation used could be that covered in the corresponding Annex III dossier (the so-called lead formulation) hence the same study could fulfil the Annex II requirement as well as the Annex III requirement. As Annex II data aim at characterising the active substance it is usually not possible to use a formulation containing additional active substances. Some lead formulations contain more than one active substance; results could be acceptable when there is no effect up to the top dose level or at the limit dose; otherwise it would be difficult to attribute the toxicity to one or the other substance.

The need for standard toxicity tests on the lead formulation (Annex III)

One Annex III package for a representative formulation has to be submitted to enable Annex I listing. Annex III contains certain study types that are also part of Annex II (standard laboratory tests with birds, bees, arthropods, earthworms and soil microorganisms). Each Annex point has to be addressed; however, it is not always necessary to generate experimental data with the formulation; instead the data on the active substance could be sufficient. The decision should be based on the following considerations:

- If the risk indicators (TER, HQ) based on the active substance are well above the TER trigger or below the HQ trigger (e.g. 100-fold) then studies with the formulation could be considered dispensable. However, a decision should be made on a case-by-case analysis in agreement with the RMS and be reported.
- It might be sufficient to test the formulation with that species of a group that was most sensitive with the active substance.
- In cases where further information is considered necessary it should be examined, whether a direct step to higher-tiered-tests would be more appropriate than repeating the basic test with the formulation.

If a notifier is of the opinion that tests with a formulation are not needed, an explanation must be given.

2.5 Endocrine effects

Endocrine disruption is to be viewed as one of the many existing modes of action of chemicals and thus can be assessed in the normal conceptual framework. However, endocrine disrupting chemicals typically affect certain life stages during reproduction and development, so potential effects may remain undetected if a test covers only a part of the reproductive cycle, as is the case in the avian one-generation study. The OECD is currently engaged in reviewing the test guidelines and where necessary improving the protocols (Task Force on Endocrine Disrupter Testing and Assessment (EDTA)). As soon as amended methodology is validated and agreed on, then this should be applied in the assessment. Meanwhile it should be considered whether evidence from mammalian studies and existing ecotoxicological studies suggests on endocrine effects such as thyroid or gonadal tumors, abnormal sex differentiation and sex organ development. In such cases the available information, e.g. from a current avian reproduction test should be re-evaluated carefully (see SCP 1999).

2.6 Higher tier tests

The data requirements (Annex III) contain a suite of higher tier tests that can be submitted if the results of the basic tests are not sufficient to decide that the risk might be acceptable and to allow for a decision with regard to inclusion of an active substance into Annex I. It should be noted, however, that (semi)field tests are not the only option for refining the assessment. Before conducting such tests other possibilities to address the problem should be considered.

Higher tier tests aim at one or more of the following purposes:

- generate information on certain parameters of the risk assessment (e.g. an avian acceptance test gives information on the palatability of potential food items which is used to refine the food consumption rate and thus the exposure estimate of the exposed species)
- investigate effects under more realistic conditions (semi-field and field tests)
- produce effects data for a wider range of species and include inter-species interactions (e.g. model ecosystems or soil community tests in the field)

Higher tier tests generally provide information on exposure and effects under more realistic conditions compared with standard laboratory tests. Therefore many uncertainties are reduced, however, as some of the variables are not under the control of the experimenter, the results tend to be less reproducible.

With regard to methods some tests such as the bee field test are standardised and fairly easily conducted. Other tests have to be planned on a case-by-case basis (e.g. terrestrial vertebrate field tests). Usually the results of the basic tests together with background information are used to define clearly the objective of the study and to select the appropriate methods, endpoints and study design in order to make sure that the study focuses on the identified concerns. Thus, the following should be considered: species at risk, type of effect (e.g. mortality or sub-lethal effects), duration of effects (e.g. are acute or long-term effects expected?), whether recovery is to be studied. When planning a higher tier study the notifier might wish to discuss the protocol with the Rapporteur Member State or consult independent experts.

2.7 Persistence

Persistent active substances and metabolites are of special concern as influences on organisms can continue to act over generations, they may have multiple effects, and any recovery may take an unduly long time. Therefore, a higher degree of scrutiny is needed to assure that non-target organisms are not affected. The assessment has to ensure that all routes of exposure are adequately considered. Persistence may be accompanied by greater bioaccumulation than would be observed for a non-persistent substance and this also should be fully addressed. Aquatic bioaccumulation data cannot be transferred to terrestrial organisms; however there are models available which describe the behaviour of an active substance/metabolite in soil organisms based on simple data (e.g. Connell and Markwell 1990, Jager 1998) as well as models to describe food chains to mammals and birds (Romijn et al. 1994). It has to be observed that not all of these models are validated, and up to now they are not routinely used for regulatory purposes. Furthermore the applicability of these models is restricted to certain chemical types.

According to Annex VI 2.5.1.1 no authorisation shall be granted "*if the active substance and, where they are of significance from the toxicological, ecotoxicological or environmental point of view, metabolites and breakdown or reaction products, after use of the plant protection product under the proposed conditions of use during tests in the field, persist in soil for more than one year (i.e. DT90 > 1 year and DT50 > 3 months), or during laboratory tests, form not extractable residues in amounts exceeding 70 % of the initial dose after 100 days with a mineralisation rate of less than 5 % in 100 days, unless it is scientifically demonstrated that under field conditions there is no accumulation in soil at such levels that unacceptable residues in succeeding crops occur and/or that unacceptable phytotoxic effects on succeeding crops occur and/or that there is an unacceptable impact on the environment, ...*"

If certain persistence triggers are exceeded, further tests with soil organisms are to be conducted (see chapter 6.1). With regard to bound residues effects on soil organisms are unlikely as long as the substance is not bioavailable. However, under certain conditions bound residues may become bioavailable and therefore a risk cannot be ruled out. Therefore it is proposed that the same data requirements should apply as for those substances with a $DT90_f$ of >365 days and a $DT50_f$ of >3 months. If there is convincing evidence from the fate data package (for example release rates, release behaviour) then further data may not be necessary.

2.8 Risk assessment

Risk characterisation

For risk assessment purposes it is common to use quotients which combine exposure and effect in order to characterise the risk. However, there are numerous ways in which such indicators could be formally defined. Unfortunately terrestrial ecotoxicology within the framework of Directive 91/414/EEC is not uniform in this regard for various reasons. Currently it uses TER values (terrestrial vertebrates, earthworms) along with HQ values (for bees). In this Guidance Document it became necessary also to introduce an indicator for arthropods taken from the ESCORT II document (Candolfi et al. 2001) where it is termed HQ. This document retains the terminology and definitions laid down in Annexes II, III and VI of 91/414/EEC. Nevertheless, it is useful to give a few explanations: Risk indicators are particular with regard to the following properties:

Direction of quotient (toxicity to exposure or exposure to toxicity)

Usually indicators under 91/414/EEC relate toxicity to exposure (TER) which means that the higher the figure the greater the safety. Exceptions are the hazard quotients (HQ) for bees and other non-target arthropods where the opposite applies, exposure being divided by the toxicity (the higher the figure the greater the risk).

Unit concordance

Mostly exposure and effects are expressed in the same unit, e.g. both as concentration in soil (mg/kg), or both as dose per body weight (mg/kg bw). This is also true for arthropods (g/ha or ml/ha). The only exception is the hazard quotient for bees where application rate (g/ha) is divided by bee LD50 (µg/bee); the latter relation makes sense, of course, as the application rate is a measure for exposure and the bee LD50 is a measure for effect. However the absolute level of the resulting HQ is meaningless without calibration; (in this case calibration has been done, see next point).

Validation, rationale for critical TER and HQ

TER values are defined such that the toxicity is taken from standard tests with the most sensitive of the tested species and the exposure is an estimate of the realistic worst case. In order to account for uncertainties (e.g. tested species vs universe of species, lab to field) assessment factors are introduced which under 91/414/EEC appear as critical TER values, e.g. 10 for the acute TER for terrestrial vertebrates and earthworms. Although founded on general experience in risk assessments the critical TERs are somewhat arbitrary (Chapman et al. 1998, SCP 2002). In contrast, the critical HQ of 50 for bees as well as the critical HQ of 2 for arthropods have a different reasoning. These values have been established according to a validation procedure where the HQ was compared with (semi)field data. The predictive power of these two HQ are therefore better defined. (It should be noted that as regards the non-target arthropod trigger value of 2, there has been some criticism due to the limited nature of the data set). Two principle points have to be observed:

- The critical HQ is only applicable to situations and conditions which have been included in the validation; for example, with both, arthropods and bees, the validation included spray applications only.
- The critical HQ is only applicable if the HQ is calculated in the same way as for validation; for example with arthropods the validation has been conducted using LR50 data from glass plate tests, not for effects data from other tests (Candolfi et al. 2001).

Interpretation of TER and HQ values

TER and HQ values should be used as indicators of risk in the assessment process. In cases where the calculated values do not meet the relevant trigger the provisions in Annex VI require that no authorization shall be granted unless it is clearly established through an appropriate risk assessment that no unacceptable effects occur under field conditions. There are several options to proceed, for example:

- refined exposure estimates
- refined effects assessment
- higher tier studies
- re-evaluation of the risk in more detail, considering the magnitude, probability and ecological significance of effects
- consideration of risk reduction measures (determined at Member State level when granting authorisations); examples are given in chapters 3.4, 4.4, 5.4, 6.4
- no authorisation of certain uses of particular concern or, finally, of all uses.

Applying risk mitigation measures and refining the toxicity and exposure estimate will result in new TER values. These amended values should be compared to the appropriate Annex VI values again to indicate whether the proposed risk mitigation measure is adequate. (HQ values underlie some constraints in this regard, see above). In higher tier studies, however, exposure is usually part of the study design, so that the results are not used for a formal TER (or HQ) calculation but immediately interpreted in terms of risk. If sufficient risk reduction measures cannot be identified, non-inclusion of the substance into Annex I of Directive 91/414/EC must finally be considered.

Example 1: The basic data may show that a product is toxic to bees with a hazard quotient clearly above the trigger of 50. If higher-tier studies confirm the risk then effective risk mitigation measures are a prerequisite for the authorisation. In this case the use could be restricted to glass-houses that are inaccessible to bees (and where no pollinators are introduced), or a label phrase could be required that would exclude applications to flowering plants (if that is compatible with the intended use of the product).

Example 2: The avian acute and dietary toxicity data for a seed treatment may indicate a high risk for seed-feeding birds with TER_a- and TER_{st}-values (according to the standard calculation) below the trigger values of 10. The refined risk assessment re-examines the worst-case assumption that birds feed exclusively on treated seed. This re-assessment is reliant upon additional data, i.e. the results of palatability studies and/or field studies. These studies may demonstrate a clear avoidance of treated seed so that it is considered unlikely that birds in the field would ingest sufficient seed to cause toxic effects and the risk may be judged as acceptable.

Probabilistic risk assessment

The traditional TER-based approach uses point estimates for the input parameters (e.g. lowest available toxicity figure, highest exposure level) and involves an overall factor (= critical TER) to cover the various sources of uncertainty. Such a deterministic assessment has limitations with regard to the quantification of the risk. This problem could be overcome by newly emerging probabilistic approaches. Performing a probabilistic risk assessment (PRA) involves assigning probability density functions to the various components that affect risk, and then carrying out Monte Carlo simulations or other calculations in order to estimate the probability that a certain event takes place. At present PRA has some shortcomings:

- For many input parameters reliable information on the distribution is lacking
- There are no common standard methods for the statistical calculations

The result of the assessment appears complex in nature and thus may be difficult to communicate to non-experts. However, that should not be regarded as a drawback.

Strengths and weaknesses of PRA methods and their applicability for regulatory purposes are presented in Hart (2001). It should be noted that some weak points such as lack of information on distributions are likewise shortcomings of current deterministic approaches. Furthermore, generic data may be used where specific data are insufficient. In conclusion, PRA methods must be regarded as promising tools and already now there may be situations where their use could be envisaged.

2.9 Metabolites

Introduction

The active substance of a plant protection product may be transformed in the environment by either abiotic or biotic processes. Under Directive 91/414/EEC, the potential risks that these metabolites pose to terrestrial organisms must be assessed.

Definitions

To facilitate clear understanding the following generic definitions are used in this guidance document:

Metabolite
For the purpose of this document, the term is used for all breakdown products of an active substance of a plant protection product, which are formed in the environment by biotic or abiotic processes after the application.

Major metabolite
All metabolites that are formed in amounts of ≥10 % of the applied amount of active substance at any timepoint evaluated during the degradation studies in the appropriate compartment under consideration.

Minor metabolite
All metabolites, degradation and reaction products that are formed in amounts of <10 % of the applied amount of substance of active substance at any time during the degradation studies under consideration.

Ecotoxicologically relevant metabolite
A metabolite which poses a higher or comparable risk to terrestrial organisms as the active substance. Such a metabolite is relevant for the overall decision on Annex I inclusion or for definition of risk mitigation measures.

Definition of ecotoxicologically significant residues (Annex VI, B.2.6.2)
An active substance or – if appropriate – a metabolite for which an analytical method has to be established for monitoring purposes (see below).

Relevant compartments

When assessing risks to terrestrial organisms, metabolites in the following media and compartments have to be considered and the potential risk for the respective organisms should be addressed:

Soil
Data on metabolites in soil come from the environmental fate section, including information on time course of appearance and concentration level. These metabolites are relevant for soil organisms and ground dwelling arthropods.

Plants
Information is provided by plant metabolism studies. Metabolites may be relevant for arthropods including bees and herbivorous birds and mammals.

Vertebrates (fish, birds, mammals)
The toxicology package contains information on absorption, distribution, metabolism and excretion in mammals. Similar data on poultry are required if, according to the intended use, residues could be found in poultry feed. In the ecotoxicological assessment, residues in vertebrates, be it the active substance or metabolites, are considered in the context of potential food chain transfer. It is not considered likely that modern plant protection products magnify in vertebrate food chains, however this route should not be ignored. Should a substance be persistent and bio-accumulative in birds, mammals or fish a proper risk assessment is necessary (for details see Appendix III of the Guidance Document on Risk Assessment for Birds and Mammals (SANCO/4145/2000)).

If exposure of a certain environmental compartment is not expected (e.g. wound-healing or stored-produce uses), further assessments are not normally required (c.f. Annex VI, C2.5.1.1, and Annex II point 7).

Requirements for assessment and testing

As a general principle, it should be understood that assessments raised in this context do not always have to be addressed by experimental studies. Notifiers are invited to address the open questions by any other available information in support of a scientific and rational assessment.

As a matter of course more supporting evidence is needed for major metabolites whereas a qualitative approach can be used for minor metabolites. Valuable sources of information include, but are not limited to:

- consideration of molecular structure of the metabolite (active part intact?);
- the occurrence of metabolites in the medium in existing tests with the active substance or major metabolites;
- with regard to birds and mammals: the appearance of the metabolite in rat and poultry (Annex points II 5.1 and II 6.2);
- general knowledge on the relationship between the toxicity of the metabolite and its parent substance (e.g. from the aquatic base set (fish, daphnia, algae);
- information on pesticidal activity from biological screening data;
- available knowledge on related compounds;
- risk indicators (TER, HQ) calculated for the parent compound (clearly on the safe side of the trigger?).

If the metabolite is CO_2 or an inorganic compound, not being or containing a heavy metal; or, if it is an organic compound of aliphatic structure, with a chain length of 4 or less, which consists only of C, H, N or O atoms and has no "structures" or functional groups which are known to be of ecotoxicological concern, then no further studies are required and the metabolite is not considered to be ecotoxicologically relevant and is of low risk to the environment.

Generally a risk assessment is needed for all metabolites. However, metabolites occurring at levels lower than 10 % (minor metabolites) only have to be considered in exceptional cases, e.g. if containing the active moiety of the molecule. By definition the PEC for a minor metabolite is lower than the PEC for the parent compound by more than a factor of 10; accordingly minor metabolites even if 10 times as toxic as their parent compound can be considered as safe, provided that the parent compound is safe and also provided that no new concern with regard to persistence is brought in. It is recognised that for technical reasons it might not be possible to identify minor metabolites. If metabolites are identified in lab studies but not in field studies then field studies should be regarded more relevant unless the difference is due to the methods applied; assessments on this should be left to enviromental fate specialists.

Tests with metabolites may not be required where they are formed relatively rapidly and are short-lived, as their toxicity may be exerted in the tests on the parent compound. This conclusion should be supported by analytical measurements or other justifiable arguments (e.g. data from laboratory or field studies). If there is more than one metabolite it may be sufficient to conduct only tests with the most important metabolite (highest amount, most comparable in structure with a.s.). If higher tier studies have been conducted with the active substance, or a relevant formulation, these studies may have also encompassed the exposure to metabolites (depending on the duration of the study and the degradation behaviour af active substance and metabolites).

Information on which tests are necessary with metabolites are found in chapters 3.1, 4.1, 5.1, and 6.1 for the different groups of organisms.

The purpose of the toxicity studies is both to establish the relative toxicity of the metabolite to the parent compound, particularly for sensitive organisms, and also to provide an effect concentration for risk assessment purposes.

Risk assessment for metabolites

In principle the risk assessment process for metabolites will be similar to that for active substances, albeit recognising that risk assessment cases will not always require specific study data for certain metabolites. If the metabolite is less toxic than the parent compound, then in most cases it does not pose greater risks than those indicated for the parent compound, so that a detailed quantitative assessment is dispensable. Exceptions are metabolites which are more persistent and bio-accumulative than the parent compound so that the long-term exposure is likely to be different.

If standard risk assessments indicate potential concerns then, as for parent molecules, risk refinement is possible either by refining effect levels or by refinement of the exposure estimate.

Defining ecotoxicological relevance

If as a result of the above risk assessment, a metabolite is considered to pose a similar or even higher risk to the terrestrial environment than its parent compound, and therefore, risk mitigation measures are needed, this metabolite is considered as 'ecotoxicologically relevant'. Such a metabolite must be included in the residue definition.

Definition of ecotoxicologically significant residues (Annex VI 2.6.2)

According to Annex VI B 2.6.2 and C 2.6.2 analytical methods must be available for post-registration control and monitoring purposes among which there are methods for residue analysis of the active substance, metabolites, breakdown or reaction products. The methods must be able to determine and confirm residues of toxicological, ecotoxicological or environmental significance. With regard to foodstuff, provisions in Annex VI contain details on sensitivity etc. With regard to environmental media, however, such specifications are missing which obviously is due to the fact that there are currently only some Member States which have maximum residue levels for soil and surface water and systematic monitoring programmes for these media. Nevertheless, definition of residues for environmental compartments is requested in the Annex I procedure. With regard to soil the following definition of "ecotoxicological significance" is proposed provisionally: Apart from the parent compound the definition should include firstly metabolites which pose a higher or comparable risk to terrestrial organisms as the active substance (= ecotoxicologically relevant metabolites according to the definition given above). Secondly, also any hazardous metabolites should be included which needs establishment of a threshold for effects data. A suitable concentration level would be that which results in the classification of a substance as environmentally hazardous. Unfortunately the EU classification system according to Directive 67/548/EEC does not yet contain criteria with regard to soil organisms, but they are in preparation. As soon as these concentrations are agreed upon they should be used for the purpose here. There is often the situation that there is no separate test with a metabolite because the metabolite appears in the system during the test with the parent compound. Then it is impossible to decide whether the observed effect is to be ascribed to the parent compound or to the metabolite. This distinction could be unimportant for the risk assessment, but the question of

whether the metabolite is hazardous remains open. In such a situation the metabolite should be regarded as ecotoxicologically significant. However, additional studies can be submitted to remove the metabolite from the residue definition. Metabolites included in the residue definition need analytical methods.

It should be noted that the definition of the residues is a formal process which is different from risk assessment.

3 Terrestrial vertebrates

3.1 Data requirements and testing

Avian acute oral toxicity (Annex II 8.1.1)

Work conducted for the UK Pesticides Safety Directorate (Hart and Thompson 1995) shows that regurgitation can substantially reduce the dose absorbed by birds in acute oral toxicity tests. Therefore, during the evaluation of avian acute oral tests it should be assessed whether regurgitation or emesis has occurred. If so, it may be appropriate to repeat the study using birds which do not regurgitate, in particular if a high risk use – such as seed treatment - is being assessed.

For example, if regurgitation is observed in an acute oral toxicity test at 500, 1000 and 2000 mg a.s./kg bw but not at 200 mg a.s./kg bw, and if there is no mortality at 200 mg a.s./kg bw then the conclusion is valid that the LD50 is >200 mg/kg bw and this figure may be used in the initial risk assessment. If this assessment raises concern, i.e. TER_a less than 10, then either an acute or dietary study will be requested using a bird species which does not regurgitate. If the initial assessment does not raise concern, i.e. TER_a >10, no further data will be requested. Sometimes regurgitation may occur in all doses whilst mortality occurs only in the top doses, i.e. regurgitation is not sufficient to protect birds. Also in this situation, a further study with a non-regurgitating species will be required.

Avian short term dietary toxicity (Annex II 8.1.2)

When the test diet has been analysed the results should be reported in the monograph. According to OECD guideline 205, a deviation up to 20 % between measured feed concentrations and nominal values is considered to be acceptable. In the case of larger deviations toxicity figures should be recalculated using effective concentrations.

Avian reproduction (Annex II 8.1.3)

A reproductive toxicity study should always be conducted unless it can be demonstrated that exposure of birds (adults and young) does not occur during the breeding season. When all relevant species are considered, the breeding season could be rather long and even short exposure periods may give rise to concern with regard to potential reproductive effects. Thus, in the case of foliar applications during the breeding season, for example, the test should normally be required even if only one treatment per season is intended.

A justification for not conducting a bird reproduction study must be supported by data to indicate that no exposure will occur during the breeding season. The justification may be

based on residue data on potential feed items. Reproductive data are always required for substances which are generally persistent (see chapter 2.7) or have a bio-accumulation potential. Reproductive data are not required, for example, if plant protection products are used indoors or if a product with a short half life of <14 days on food items is applied in autumn. It should be noted that low acute and dietary avian toxicity are not sufficient to indicate a low reproductive toxicity.

Effects of secondary poisoning (Annex III 10.1.4)

Annex point III 10.1.4 mainly addresses the food chain from rodents to predators and scavengers in the case of rodenticides. For further information see Doc SANCO/4145/2000.

Metabolite testing

Metabolites in or on potential feed items have to be considered. However, apart from general considerations explained in chapter 2.9, there are some cases where experimental toxicity testing is not necessary:

- If the metabolite in question also appears in birds and mammals it can be assumed that any toxic effects would be expressed in the toxicity test with the parent compound, and that the risk from the metabolite is covered. It has to be observed that the toxicology section of the dossier/monograph always provides information on metabolism in rats, but not necessarily on metabolism in birds (poultry), and it cannot be assumed that the metabolic pathway in birds is identical to that of mammals.
- The toxicology data package may already contain mammalian toxicity tests with the metabolite. The absolute toxicity of the metabolite cannot be directly extrapolated from mammals to birds, but the relation can be used as an indication that such information might be sufficient for an assessment. For example, consider the following data and information:
  LD50 rat (parent) = 238 mg/kg,
  LD50 rat (metabolite) = 680 mg/kg,
  LD50 quail (parent) = 42 mg/kg.
  So, in rats the metabolite is 2.9 times less toxic than the parent. One should refrain from multiplying the quail LD50 (parent) by 2.9 because that would imply an undue level of accuracy. However, it would be reasonable in most cases to assume that also in birds the metabolite is not more toxic than the parent compound.

Should testing become necessary an acute oral study would be the first choice to serve as a bridging study, i.e. to compare the inherent toxicity of the metabolite with that of the parent compound.

3.2 Exposure assessment

Exposure assessment is dealt with in Doc SANCO/4145/2000

3.3 Risk assessment

Risk assessment is mainly dealt with in Doc SANCO/4145/2000. Therefore this chapter only contains some additional information.

Relevant toxicity figure for the acute assessment

Calculation of TER_a should be determined using the lowest, reliable acute oral LD50 figure. If data on the acute toxicity of both active substance and formulation are available, it should be determined whether animals are likely to be exposed to the formulation or the active substance and the more appropriate figure should be used. For instance, in the case of granules birds are clearly exposed to the formulation whereas in the case of a spray application, residues on green plant material are better considered in terms of the active substance than of the formulation.

3.4 Risk mitigation options

Risk mitigation is dealt with in Doc SANCO/4145/2000.

4 Bees

For general background information see the upcoming EPPO scheme (EPPO 2002b)

4.1 Data requirements and testing

Acute toxicity to bees (Annex II 8.3.1.1, Annex III 10.4.1)

If honeybees are likely to be exposed to the active substance both acute oral and contact toxicity tests must be conducted as the toxicity by one route of exposure cannot be predicted from the other. Where there is only one relevant route of exposure (e.g. oral exposure in the case of soil application), testing can be restricted to this exposure route. The test result should be presented as µg a.s./bee or µg formulation/bee. If there are problems with solubility of the active substance, then the test should be conducted with a representative formulation.

Toxicity tests should be conducted according to EPPO 170, or OECD 213 and OECD 214 guidelines.

Bee brood feeding test (Annex II 8.3.1.2)

The test method of Oomen et al. (1992), that is recommended in Annex II for insect growth regulators, is a worst case screening test. If no effects are found the conclusion is justified that no brood damage will occur when using the product. In the case of effects further cage/tent/tunnel or field studies are necessary to evaluate the risk under more realistic conditions. If toxicity to honeybee broods can already be predicted from the mode of action of the compound, testing may immediately start with cage/tent/tunnel or field trials.

Residue test (Annex III 10.4.2)

Aged residue tests may be valuable as an additional tool for risk assessment. However, no specific validated methods are yet available. The test should be designed to assess the duration of effects due to residual traces of plant protection products on the crop.

Higher tier tests (Annex III 10.4.3, 10.4.4 and 10.4.5)

For higher tier testing (cage/tent/tunnel or field trials), the recommendations of EPPO guideline 170 should be taken into account.

Testing of systemic plant protection products

For soil-applied systemic plant protection products (e.g. plant protection products applied as seed dressing) the acute oral toxicity of the active substance(s) have to be determined. If potential risks to honeybees are identified (i.e. very low LD50) realistic exposure conditions should be taken into account, i.e. realistic exposure concentrations as expected in nectar and pollen as indicated by residue studies. If a risk is indicated, higher tier studies (cage/tent/tunnel or field studies) with realistic exposure scenarios should be performed.

Metabolite testing

Standard lab tests are normally not required for metabolites. Exceptions may be cases where for example the metabolite is the pesticidal active molecule. Before conducting studies the general guidance given in chapter 2.9 should be observed. If higher tier studies (cage/tent/tunnel or field) are conducted with the plant protection products under realistic exposure conditions, potential risks from metabolites should be covered.

4.2 Exposure assessment

For products applied as sprays where risk as assessed according to the HQ approach exposure should be established as the maximum single application rate of the product expressed as g/ha because the HQ was validated on this measure.

For systemic plant protection products, exposure considerations and calculations should be based on the a.s. (or metabolite) present in the respective plant parts (e.g. nectar, pollen) to which honeybees could be exposed. However, it should be noted that estimates of these concentrations are rarely available.

Exposure calculations in higher tier studies are already considered within the experimental design (e.g. honeybees foraging on treated field crops).

4.3 Risk assessment

Hazard quotient for bees (Annex III 10.4)

The hazard quotient is stated to be application rate/oral LD50 or application rate/contact LD50, where the LD50 is expressed as µg a.s./bee and the application rate is in g a.s./ha. As stated above, the maximum single application rate should be used to calculate the oral and contact HQ-values. If the oral and contact HQ < 50, low risk to bees is concluded and no further testing is required. If the oral or contact HQ > 50, further higher tier testing is required to evaluate the risk to bees. The critical HQ of 50 was validated against incidents (EPPO 2002b); it is only applicable to spray products.

Higher tier risk assessment for bees

There are no clearly defined endpoints for higher tier studies, therefore, a degree of expert judgement is required to interpret both semi-field and field study results. As regards semi-field trials, where there are replicated studies, there should be a statistical comparison between key parameters, e.g. foraging density, mortality, proportion of adults, larvae and pupae in the hive. It should be noted that the parameters considered should be relevant to the compound under

consideration. For example if an insect growth regulator was being assessed then it would be more relevant to concentrate on developmental issues. As regards field trials, key parameters should be compared to either pretreatment levels or to control levels. It is important to consider any effects observed in relation to the overall survival and productivity of the hive. Key parameters which may be considered in a field trial include: mortality (assessed via the use of dead bee traps), behaviour (including foraging behaviour in the crop and around the hive), honey crop (assessed via weighing the hive at appropriate intervals) and state of colony (including an assessment of brood). Depending upon the concern highlighted in the initial risk assessment it may be appropriate to use pollen traps as well as appropriate analysis of dead bees. Analysis of honey and wax may be useful in determining exposure. The use of a toxic standard in both semi-field and field trials along with an untreated control can aid interpretation of the results. For insect growth regulators and other active substances which may cause long-term adverse effects on hive health, evidence is required confirming a lack of effects on hive health over a long time period. It should be noted that further information is available in the EPPO guideline (EPPO 2001). The design of higher tier studies is dependant upon the risks highlighted and therefore it is recommended that applicants should consult the relevant authority.

4.4 Risk mitigation options

The risk mitigation measures outlined below are options only. These measures will require consideration at a national level and implementation will depend on local agronomic practice and conditions. If predicted effects to honeybees are considered as not acceptable, the following aspects of the use pattern may be considered for modification in order to mitigate the predicted risk:

- application rate
- timing of application (e.g. apply in the evening after honeybee flight, do not apply during honeybee flight)
- GAP adaptation (e.g. do not apply during crop flowering)
- agronomic practice (e.g. mulch ground cover before application of the plant protection products)

5 Other arthropods

The risk to non-target arthropods is routinely assessed under 91/414/EEC. Annex II of 91/414/EEC states that data on two sensitive standard species as well as data on two crop relevant species are required. If effects are observed with species relevant to the proposed use then further testing may be required. Annex III of 91/414/EEC states that where significant effects have been observed the toxicity of the product to two additional species must be investigated. Both Annex II and III reference the SETAC Guidance document on regulatory testing procedures for pesticides with non-target arthropods (ESCORT, Barrett et al. 1994) as a source of guidance for testing. However, several limitations have been identified and these can be summarised as:

- The objectives of the testing scheme are not clear, e.g. it does not precisely discriminate between non-target arthropods in a general context and beneficial arthropods in an agricultural or IPM context.

- The trigger value for first tier data (30 % effects as laid down in Annex VI C point 2.5.2.4) leads to excessive higher tier testing.
- The single-dose laboratory data generated do not provide for determination of the intrinsic toxicity of the substance (except where is no effect and the test can be regarded as a limit test). In addition this kind of testing is inflexible and does not allow a satisfactory risk assessment especially for off-field habitats.
- Uncertainty about data requirements, testing methodology and evaluation, especially for multiple application products, where currently life span, spraying interval and fate are ignored and for off-crop habitats, where exposure scenarios and mitigation measures are not yet agreed.

Due to the above issues a workshop, ESCORT 2, was held in 2000 which aimed to address these shortcomings. From this workshop a guidance document resulted (Candolfi et al. 2001) which is referred to here as "ESCORT 2". This workshop was attended by all EU Member States as well as representatives from industry and academia and revised the process by which the risk to non-target arthropods should be assessed. By building on the experience gained from assessing the risk to non-target arthropods under 91/414/EEC, a new approach was proposed which offers a high level of protection, but is more focused and structured.

The process discussed and agreed on this workshop starts with glass-plate tests on the two standard sensitive species referred to in Annex II (*Aphidius rhopalosiphi* and *Typhlodromus pyri*). However, rather than a single rate study, a rate-response study is usually required. The endpoint of these studies are LR50 values (i.e. lethal rate that causes 50 % mortality) which are compared to the predicted exposure both in-field and off-field. With substances suspected to have a special mode of action (IGRs, insect feeding inhibitors) tests should include sublethal endpoints and may need other modifications. The assessment of risk for arthropods living in- and off-field is conducted separately. If the resulting 'hazard quotient' (HQ) based on the standard tests is greater than or equal to 2 then further data and/or risk management measures are required. Note: The critical trigger of 2 was proposed on the basis of the available data. It was noted at the ESCORT 2 workshop that this value should be revised when suitable data are available.

It is proposed that for active substances and their associated product(s) under consideration for inclusion on Annex I, the risk to non-target arthropods both in and off-field should be adequately addressed. The guidance given below is in line with the recommendations of ESCORT 2.

5.1 Data requirements and testing

Standard tests (Annex II 8.3.2, Annex III 10.5.1)

Testing is always required where exposure of non-target arthropods is possible.

Standard tier 1 testing comprises glass plate tests with *Aphidius rhopalosiphi* and *Typhlodromus pyri*. Preferably these tests should be designed as rate-response studies in order to determine the LR50 as this allows for applying the data to different use scenarios and also to the risk assessment for off-crop areas. However, if the toxicity is expected to be low then limit tests can be conducted at a rate equivalent to the maximum application rate multiplied by the multiple application factor (MAF). With regard to the test substance (active substance,

lead formulation) see chapter 2.4. With substances suspected to have a special mode of action (e.g. IGRs, insect feeding inhibitors) tests should include sublethal endpoints and may need other modifications.

Details on methods are given in the ESCORT 2 document.

Higher-tier tests (Annex III 10.5.1 and 10.5.2)

Higher-tier tests are required when a risk is indicated in lower assessment tiers. There are several options for higher-tier testing or combinations of adequate tests:

- Extended laboratory tests (tests with natural substrate aiming at lethal and sublethal effects)
- Aged-residue studies
- Semi-field tests
- Field tests

ESCORT 2 provides advice regarding the choice of studies and the selection and number of species. Usually these studies are conducted with one dose rate matching the field application rate taking into account multiple applications and the use of appropriate risk mitigation measures. Advice is given in ESCORT 2 regarding the appropriate rates to use in such studies. With regard to extended laboratory tests it should be noted that due to the implementation of a correction factor[1] (default value = 5) in some cases the rules may give application rates greater than the field rate. In this case it is suggested to test at the maximum rate including the multiple application calculation. In the case of extended laboratory studies a dose response design may be more informative than a one-dose design.

Metabolite testing

Arthropods may be exposed to metabolites in/on plants and to soil metabolites.

Metabolites in vegetation: Standard lab tests are normally not required for metabolites. Exceptions may be cases where for example the metabolite is the pesticidal active molecule. Before conducting studies the general guidance given in chapter 2.9 should be observed. If higher tier studies (semifield or field) were conducted with the plant protection products under realistic exposure conditions, potential risks from metabolites should be covered.

Soil metabolites: These are assessed with regards to soil organisms, so that tests with soil-surface arthropods are not needed.

[1] In order to avoid confusion the terminology of the ESCORT document is used in this document as far as possible; actually "uncertainty factor" or "safety factor" would be more appropriate

5.2 Exposure assessment

Generally, exposure for non-target arthropods is expressed in terms of application rate (g/ha or ml/ha).

Tier I assessment

For the standard assessment the following scenarios are used to describe the exposure in-field and off-field. For both, the key input is the nominal field application rate supplemented by various factors:

in-field exposure = Application rate * MAF

off-field exposure = Application rate * MAF * (drift factor / vegetation distribution factor)

For calculation of MAF values, definitions and further details see ESCORT 2. With regard to the vegetation distribution factor ESCORT 2 gives a default value of 10. However, this figure is considered unreliable, therefore more appropriate data should be used as soon as they become available (a research project is currently under way). With regard to the drift factor the tables published by Rautmann et al. (2001) may be used; the standard assessment should be conducted for 1 m distance (arable crops) or 3 m (orchards and vineyards); drift factor = % drift / 100.

Basic drift values for one application Ground deposition in % of the application rate (90th percentiles)									
Distance	Field crops	Fruit crops		Grapevine		Hops	Vegetables Ornamentals Small fruit		Field crops
[m]		Early	late	Early	late		Height < 50 cm	Height > 50 cm	Water > 900 l/ha
1	2.77						2.77		4.44
3		29.20	15.73	2.70	8.02	19.33		8.02	
5	0.57	19.89	8.41	1.18	3.62	11.57	0.57	3.62	0.18
10	0.29	11.81	3.60	0.39	1.23	5.77	0.29	1.23	0.05

Higher-tier assessments

Refined assessments are based on the outcome of higher-tier studies. In such studies relevant exposure issues are considered in the study when establishing the dosing regime (be it dose-response design or single-dose design). That makes a separate exposure assessment unnecessary; it must, of course, be ensured that the study covers the use scenario under assessment.

5.3 Risk assessment

Assessing the risk 'in-field'

Step 1: Tier I assessment based on standard tests

In the first tier the risk is characterised by the 'in-field' hazard quotient (HQ):

In-field HQ = in-field exposure / LR50

where the LR50 comes from glass-plate tests with the two standard species. If the in-field HQ is less than 2 for both species, no further assessment is required (for the reasoning behind this trigger level see ESCORT 2). If the HQ is greater than or equal to 2 for one or both species then go to step 2.

Step 2: Higher tier assessment

If no appropriate risk mitigation measures can be identified, then the notifier should carry out higher tier studies on the affected species and one further species with different biology. Details of suitable species are provided in ESCORT 2. With regard to extended laboratory tests and semi-field tests lethal, and sublethal effects of less than 50 % are considered acceptable provided that the tests covered the appropriate field rate. For interpretation of aged residue studies with respect to recolonisation, and for interpretation of field studies see ESCORT 2. Generally, it has to be demonstrated that there is a potential for recolonisation / recovery at least within one year but preferably in a shorter period depending on the biology (seasonal pattern) of the species. The assessment may be based on field studies or other evidence (e.g. results of aged-residue studies, environmental fate information). In any case the data and assumptions should be fully justified.

Assessing the risk 'off-field'

Step 1: Tier I assessment based on standard tests

The risk is characterised by the 'off-field' HQ:

Off-field HQ = (off-field exposure / LR50) * correction factor

where the LR50 comes from glass-plate tests with the two standard species; the correction factor is intended to cover uncertainty with regard to species sensitivity, the default value is 10. If the off-field HQ is less than 2 for both species, no further assessment is required, if greater than or equal to 2 for one or both species then go to step 2.

Step 2: Higher tier assessment

If no appropriate risk mitigation measures can be identified, then higher-tier studies on the affected species and two additional species with different biologies should be conducted. Details regarding suitable species are provided in ESCORT 2. With regard to extended laboratory tests and semi-field tests lethal and sublethal effects of less than 50 % are considered acceptable provided that the tests covered the appropriate field rate; the default value for the correction factor is 5. Generally, it has to be demonstrated that there is an acceptable potential for recovery within an ecologically relevant period.

Basically, if the tier-1 assessment indicates a risk either risk mitigation measures or higher-tier studies are called for. It should be noted that in order to achieve Annex I listing that it is not

considered appropriate to propose unrealistic risk mitigation measures (e.g. exaggerated buffer zones) in order to avoid higher-tier testing.

Risk from solid formulations, products with a special mode of action and those of limited solubility

The standard approach is not appropriate for substances with limited solubility or for plant protection products such as granules, seed treatments and pellets. In these cases it is recommended that studies are conducted with *Hypoaspis aculeifer* or *Folsomia candida* as proposed by EPPO (2002a). If deemed appropriate, studies with *Aleochara sp.* might be conducted, e.g. at tier 2.

It is recognised that the standard approach may not be wholly appropriate for insect growth regulators or other compounds with particular modes of action. For these compounds the principles of ESCORT 2 should be followed with case-by-case modification according the specific issues for the compound in question.

5.4 Risk mitigation options

In order to reduce effects on non-target arthropods within the cropped area the following use specifications may be modified:

- application frequency and intervals
- timing of application (crop stage)
- unsprayed headlands

In order to reduce effects in off-field areas there are the following options:

- Buffer zones
- Wind breaks
- Drift-reducing application techniques

For further explanations see ESCORT 2

6 Soil organisms

6.1 Data requirements and testing

Acute effects on earthworms (Annex II 8.4, Annex III 10.6.1.1)

Testing is always required where contamination of the soil is possible. With regard to the test substance (active substance, lead formulation) see chapter 2.4.

Tests according to OECD Guideline 207 and ISO 11268-1: 1993 (which are similar to 88/302 EC) are also acceptable.

Sublethal effects on earthworms (Annex II 8.4.2, Annex III 10.6.1.2)

According to Annex II the requirement for this test depends on the exposure pattern to the active substance ('continued or repeated exposure'). The following triggers for persistence of the active substance and the number of applications are proposed:

- The test is not required when both the $DT90_f$ is less than 100 days, and the number of applications is less than 3.
- The test is always required if the $DT90_f$ is above 365 days (regardless of the number of applications).
- The test is always required if the number of applications is greater than 6 (regardless of persistence).
- If the $DT90_f$ is between 100 and 365 days and/or the number of applications is between 3 and 6, a case by case decision is made.

With regard to substances forming bound residues see chapter 2.7.

The test is also required if the assessment of the acute risk gives a TER of less than 10 (see below).

Suitable methods are ISO 11268-2:1998 and the forthcoming OECD 222. With products intended to be sprayed, surface application should be preferred (annex D of the ISO guideline) and the result given in g/ha. The test should preferably be conducted as dose-response test.

When planning the test, the upper concentration level must be chosen to be high enough in order to be able to judge whether the long-term TER meets the trigger of 5, which is provided in Annex VI of Directive 91/414/EC. It has to be taken into account that exposure under field conditions may be elevated due to repeated applications (see chapter 6.2) and that toxicity figures may be corrected for f_{oc}. If available and appropriate, data from field dissipation studies should be considered.

Earthworm field studies (Annex III 10.6.1.3)

The study is required where TER_{lt} is < 5. However, as already explained in chapter 2.6 it should be checked in such cases whether there are other options for refinement (EPPO 2002a).

The study should reflect the use of the compound, the environmental conditions and species that will be exposed. If the chemical is to be applied in the arable situation it should preferably be applied to bare soil as opposed to grassland where it may become bound to the surface thatch. Analysis of the soil would assist in confirming whether the field study is appropriate for the intended arable crop use. With regard to the dosage the test should be designed such that the highest exposure according to the intended use of the product is covered. That means that multiple applications should be made where relevant, and crop interception should be considered. If accumulation in soil is expected then a rate equivalent to the long-term (pluriannual) plateau concentration should be added. The type of application of the test substance (surface application, incorporation, etc.) should be according to the intended use.

A method is described by ISO (11268-3:1999). For further information see also Greig-Smith et al. (1992) and Sheppard et al. (1997). General remarks on higher tier tests (chapter 2.6) should be observed.

Soil nitrification and carbon mineralisation (Annex II 8.5, Annex III 10.7)

Testing is always required where contamination of the soil is possible.

With regard to methods, Annex III of Directive 91/414/EEC refers to a SETAC document (Lynch 1995). In the interim, the OECD has published its guidelines 216/217 which should be preferred when conducting new studies.

Other soil non-target macro-organisms (Annex III 10.6.2)

This Annex point requires additional data for soil organisms contributing to organic matter breakdown, depending on active substance degradation rate and on available information with regard to effects to various organisms. Principally the risk to this group of organisms, which include soil mesofauna and macrofauna, could be determined either at a species level or at a functional level. While a candidate test for the former would be a Collembola reproduction test or a test on gamasid soil mites, a candidate for the latter would be the "litter bag" test.

This Annex point particularly deals with the problem of persistent active substances or persistent metabolites ($DT90_f$ > 100 days). These are of special concern as influences on organisms can continue to act over generations and may have multiple effects, and any recovery could take an unduly long time. Therefore, a higher degree of scrutiny is needed to assure that soil organisms are not affected.

Based on the recommendations of the Lisbon Workshop (EPFES 2002) the following tiered procedure is proposed (see figure 1):

a) Collembola reproduction test or test on gamasid mites

Testing is required where contamination of soil is possible and $DT90_f$ is between 100 and 365 days and the standard HQ for arthropods (*Typhlodromus* and *Aphidius*) >2. This test is used as a potential waiver for the litter-bag-test (see next point); so, if the litter-bag test is triggered anyway by other criteria (effect on soil micro-organisms >25 % or TER_{lt} for earthworms <5) then this test could be omitted. A suitable protocol for the Collembola test is the ISO method 11267:1999; a test design with the gamasid mite *Hypoaspis aculeifer* is described by Løkke and Van Gestel (1998) and Bakker et al. (2002). As long as these methods are not validated protocols should be checked with the Rapporteur Member State.

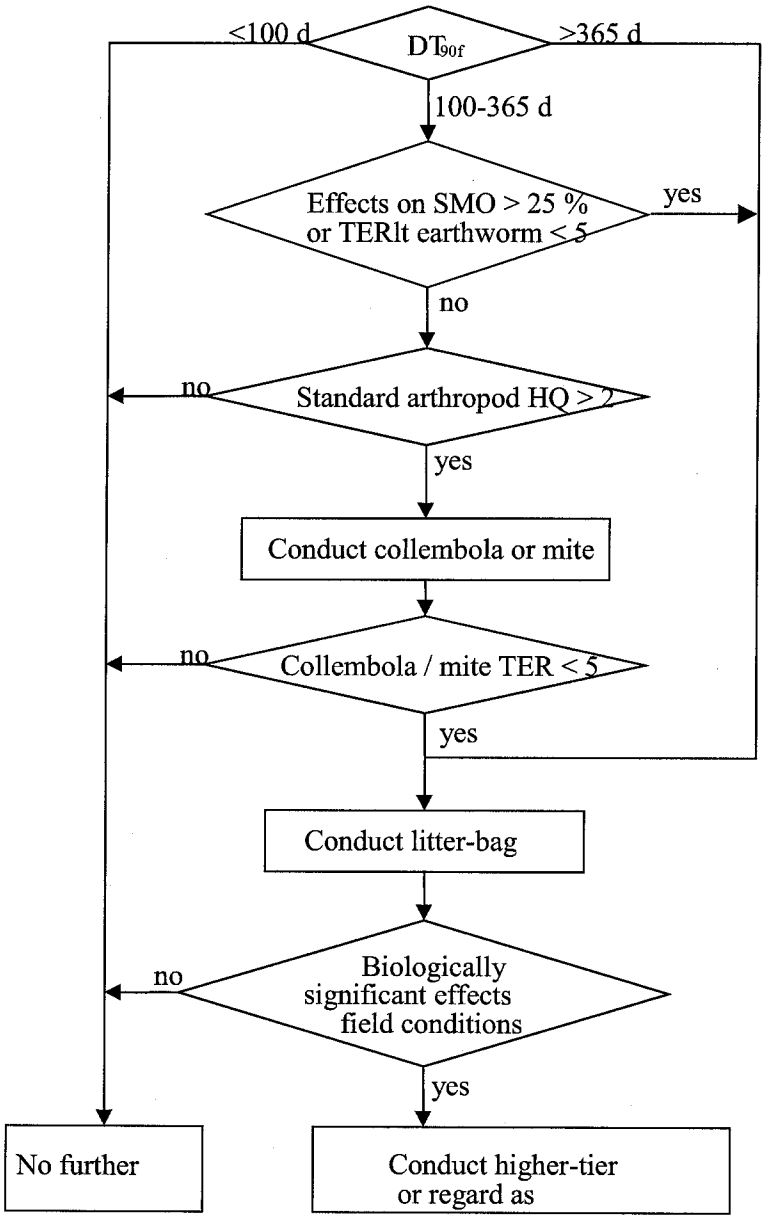
b) Litter bag test under field conditions

Testing is always required where contamination of soil is possible and DT90 is > 365 days or mineralisation is < 5% in conjunction with bound residue formation of > 70%. Testing is conditional where $DT90_f$ is between 100 and 365 days; in such cases the following auxiliary criteria are applied:
- Effects on soil microorganisms >25 % after 100 d
- or long-term TER for earthworm < 5
- or TER for Collembola or soil mites < 5

Principally this means that in the intermediate persistence range a litter bag test is not required if the above mentioned groups of organisms pass the standard tier 1 assessment.

As regards methods the test should be conducted in the field on arable sites, taking into account the intended use of the product. Concerning exposure, the plateau concentration

Figure 1: Test sequence with regard to soil organisms for persistent substances



should be applied to the soil or already be available in the soil, before the litter bags are buried. (Plateau concentration refers to the long-term pluriannual plateau over years (FOCUS 1996)). After that the annual rate is applied considering the crop interception. The degradation of fresh incorporated organic material is evaluated using at least 3 sampling dates. Minimum duration of the test should be 6 months. Special attention should be given to the method of application and the number of time points for measurement. Weight loss and the degradation rate of the organic material are the endpoints of the test. A method has been drafted at the Lisbon workshop which will appear in the workshop proceedings (EPFES 2002). As long as there are no formally harmonised protocols a certain degree of flexibility must be conceded. So, when judging the acceptability of a study it should be considered what the state of technique was when the study had been generated.

c) Higher tier tests

If the litter bag test shows biologically significant effects or there is other reason for additional concern then further testing could be an option; (there are other options such as risk mitigation; there also could be the final conclusion that there are no safe uses). If further testing is envisaged then it should be decided on a case by case basis which approach is most helpful:

- extend the on-going litter bag study or start a new litter bag study under more realistic conditions (the study may be extended for mesofauna structural endpoints; see for example Elkins and Whitford (1982), Bjørnlund et al. (2000), van Vliet et al. (2000)).
- large-scale field studies
- terrestrial model ecosystems

In any case problems and questions with the substance should be identified prior testing and tests then be targetted to these problems.

Metabolite testing

With regard to metabolite testing see general remarks in chapter 2.9. If testing of soil metabolites on soil organisms is necessary the first step should be an acute toxicity study with earthworms to compare the inherent toxicity with that of the parent compound. A particular situation may arise when the metabolite is more persistent than the parent compound. Certain tests with soil organisms are triggered by persistence (earthworm reproduction test, litter bag test, etc.), and it is possible that the persistence of the parent compound does not exceed the trigger for these studies, but the metabolite does. In such cases the additional studies should be conducted, with the metabolite, regardless of its acute toxicity.

6.2 Exposure assessment

Earthworms

The exposure is represented by the predicted in-field concentration of the substance in soil. PEC values for the various use scenarios are supplied by the environmental fate section. Initial PEC values are decisive in this context (no time-weighted averages). In the case of repeated applications, the PEC after the last application is relevant. In case of persistent substances the plateau concentration is relevant.

Soil micro-organisms and other functional tests

No separate exposure assessment is necessary for soil micro-organisms as the relevant exposure conditions (multiple application, etc.) are considered in establishing the dosing regime for the test. So the outcome of the study is immediately interpreted in terms of risk. The same is true for litter bag tests.

6.3 Risk assessment

Standard risk assessment for earthworms

The standard risk assessment is based on TER values. The acute TER is the ratio between the LC50 from the acute test and the PEC. The long-term TER is the ratio between the NOEC from the reproduction test and the PEC.

Both acute and reproductive tests are static tests where the test substance is applied to the system only once at the beginning. Therefore, the nominal dose levels in the test match initial concentrations in the field and thus it is appropriate to use initial PEC values (no time-weighted averages) for the acute as well as for the long-term TER. If it can be demonstrated that degradation in the artificial substrate and natural soils differ significantly, then it may be considered in the assessment.

The toxicity of lipophilic organic contaminants to soil organisms usually depends on the organic carbon content (f_{oc}) of the substrate as this governs adsorption and thus pore water concentration. The artificial substrate of the earthworm laboratory tests has a higher f_{oc} than many natural soils, so it could be expected that the LC50 or NOEC would be lower if the test were conducted in natural soil (Van Gestel 1992). The risk assessment should account for this difference by dividing the LC50 and the NOEC by 2 where $\log K_{ow}$ is greater than 2 (EPPO 2002a) unless it can be demonstrated by soil sorption data or other evidence that the toxicity is independent of f_{oc}. For sake of clarity the corrected toxicity figures should be denoted by a subscript (e.g. $LC50_{corr}$).

Refined risk assessment for earthworms

If the acute TER is below 10 or the long-term TER is below 5 further action is required. For general considerations see chapter 2.8. It should be decided on a case-by-case basis which option is best suited to proceed. Refinement of exposure, for example, is often quick and inexpensive and should be considered first before turning to higher tier tests.

Refined effects assessment

When the NOEC from the reproductive test is expressed in g/ha, it could be converted into mg/kg soil by a calculation assuming 100 % of substance reaching the soil, 5 cm depth and a soil density of 1.5 to give a value used in the TER_{lt} calculation. When the TER_{lt} is close to the trigger value, the calculation could be refined by considering actual test values (application rate and surface of the test unit, dry soil weight in the test unit). If there are uncertainties arising from the fact that the standard tests are conducted with artificial soil then an option might be to do the earthworm test in natural soil.

Refined exposure estimate

The exposure assessment could be improved, for example, by employing more sophisticated models, consideration of interception, or inclusion of field measurements.

Higher tier studies

Where the acute TER does not meet the trigger the earthworm reproduction test can be regarded as the next higher tier. (Note: The earthworm reproduction test fulfils two purposes. Firstly, it is a long-term test with sublethal endpoints which has its own place in the base set and is triggered by exposure considerations (continued, repeated). Secondly, it can be regarded as a higher-tier test above the acute test because it involves more realistic conditions (surface application instead of mixing into the soil).

Risk assessment for soil micro-organisms

The outcome of the soil micro-organism test is directly assessed in terms of risk. The decisive parameter is the magnitude of effect compared to the untreated control (be it increase or decrease of activity), and the time-course of recovery. According to Annex VI of 91/414/EEC the critical level is $\pm$25 % after 100 days. Larger deviations will require refinement of the assessment. As a matter of course, the concentrations used in the test must cover the maximum PEC. Generally the test concentrations are converted by calculation to equivalent doses in g/ha. Different modes of calculations are used and thus may introduce a bias in the interpretation of the risk. It is recommended to compare directly the test concentrations to the PEC values before to conclude on potential risk.

Risk assessment for non-target mesofauna

Data from a Collembola reproduction test or a soil mite test could be treated in a risk assessment in the same way as data on earthworm reproduction (TER values using PEC and NOEC)

6.4 Risk management options

Risk mitigation options for soil organisms are limited. There are possibilities to reduce the exposure (reduction of application rate and/or number of applications and/or restriction on glasshouse use only), but inevitably these measures will compromise the agricultural objectives.

7 Non-target plants

The risk of plant protection products to terrestrial plants has been until now included in a generic assessment on 'other non-target organisms (flora and fauna) believed to be at risk.' However, this aspect is considered a critical element in the evaluation of certain plant protection products, particularly herbicides and plant growth regulators, and therefore some general guidance is included.

A key element in the evaluation is the definition of non-target plants. For a generic evaluation, as required by Directive 91/414/EC, the following working definition is suggested: Non-target plants are non-crop plants located outside the treatment area.

7.1 Data requirements and testing

Annex II and III of Directive 91/414/EEC do not contain specific data requirements for non target plants. However, the introductions to these annexes generally state that there is a need to report all potentially adverse effects and to undertake additional studies where there are indications of such effects. Therefore a tiered approach is suggested starting with available data and proceeding to further steps in case of need. Data are not required, where exposure is negligible, e.g. in the case of rodenticides, substances used for wound protection or seed treatment, or in the case of substances used in stored products or in glasshouses.

Tier 1: Initial screening data

For the first tier, a preliminary assessment is conducted using available information. Preference is given to screening data; there should be at least 6 species from different taxa tested at the highest nominal application rate (1 x). These data could be supplemented by further information on efficacy, selectivity, phytotoxicity, etc. included in the biological dossier or obtained from the different field assays such as efficacy trials, residue studies, environmental fate and ecotoxicological studies, etc. The initial step is unprofitable for herbicides and plant growth regulators as these inevitably will end up in the second tier.

Tier 2: Bioassays on terrestrial plants

If a potential risk is identified (more than 50 % effect for one or more species at the maximum application rate, see chapter 7.3), then specific information on the toxicity of the substance to terrestrial plants should be requested. The second tier considers laboratory assays on a selection of plant species. It is recommended to conduct dose-response tests on 6-10 plant species representing as many taxonomic groups as possible. In order to generate data that are useful for probabilistic approaches there should not be a focus exclusively on species assumed to be the most sensitive. If, from the screening data, a specific mode of action is evident, or strong differences in the species sensitivities are identified, this evidence should be used in the selection of the appropriate test species. This may be especially true if non-herbicides reach tier-2 testing.

For foliar applications, the bioassays should be conducted by spraying the product on the plants, reproduce as far as possible the realistic exposure conditions and, in particular, spray drift. Soil application should be chosen if that is more appropriate with regard to the mode of action. The test substance should be the lead formulation (or another formulation) because formulations contain, besides the active substance, all those components and co-adjuvants required for maximising biological activity. For systemic products applied on the ground/soil, the tests should reproduce this application pattern.

Suitable test methods are the new draft OECD Guideline 208 and the OPPTS guidelines of the US EPA.

Tier 3: Field or semi-field studies

The third tier requires semi-field or field assays, to study the effects observed on non-target plants during realistic applications. Such studies are time-consuming and expensive; before undertaking them it should be checked whether there are options for the refinement of exposure and/or effects. Furthermore, as for all other non-target organisms, field or semi-field

studies are not required if the risk based on the tier 2 assessment could be managed by risk mitigation measures which could be dealt with on a Member State level.

Field or semi-field studies with non-target plants are not standardised. Therefore notifiers might wish to discuss the protocol with the Rapporteur Member State. Generally, effects on plant abundance and biomass production at different distances from the crop or at exposure levels representing different distances from the crop should be analysed. These studies are compatible with most semi-field and field studies.

7.2 Exposure assessment

Spray drift is considered the key exposure route for terrestrial plants located in the vicinity of the treated area. The drift models produced by the BBA for the exposure assessment of aquatic organisms may be used as a surrogate to cover the exposure assessment of terrestrial plants (Ganzelmeier et al. 1995, recently updated by Rautmann et al. 2001). The following table shows the drift expressed as percentage of the applied dose:

Basic drift values for one application Ground deposition in % of the application rate (90th percentiles)									
Distance	Field crops	Fruit crops		Grapevine		Hops	Vegetables Ornamentals Small fruit		Field crops
[m]		Early	late	Early	late		Height < 50 cm	Height > 50 cm	Water > 900 l/ha
1	2.77						2.77		4.44
3		29.20	15.73	2.70	8.02	19.33		8.02	
5	0.57	19.89	8.41	1.18	3.62	11.57	0.57	3.62	0.18
10	0.29	11.81	3.60	0.39	1.23	5.77	0.29	1.23	0.05

In fruit, grapevine and hops for herbicides (but not for plant growth regulators) that are applied to the ground, the column "field crops" is applicable.

It should be noted that these drift data have been generated with regard to intake into surface waters. In particular, there is no vegetational barrier between the spray boom and the collector plates. In terrestrial scenarios, however, horizontal and vertical interception by in-crop or off-crop vegetation as well as patchy distribution is relevant ("three-dimensional-situation"); thus, when more realistic drift data become available they should be used.

The initial assessment should be conducted for a distance of 1 m from the field edge for field crops, vegetables or ground applications such as for herbicides, and 3 m for other crops. Risk mitigation measures based on buffer zones within the crop area can also be quantified using the above table. In case of aerial applications a deposition rate of 100 % is assumed as the default, however this figure may be refined by applying appropriate models (e.g. AgDrift).

7.3 Risk assessment

A tiered approach with three different steps is also recommended.

Tier 1: Initial decision on the likelihood for terrestrial plant effects

This assessment step is based on the information described above as "initial screening data". The endpoints measured in most screening studies, such as phytotoxicity, chlorosis, etc. cannot be interpreted as a NOEC value covering germination and biomass production. However, the available information usually allows the use of a conservative approach, assuming, for example, that when an untreated control has been run in parallel, any effect accounting for at least 50 % reduction in biomass production could be identified in a visual inspection. In addition, single dose experiments reported in terms of percentage of observed effects can also provide indications on the potential hazard of the substance for terrestrial plants.

The detection of potentially sensitive species in this initial assessment, or the evidence of specific mechanisms of action suggesting effects on terrestrial plants (which is evident in the case of herbicides) will trigger the need for a proper quantitative assessment. As a general rule, the risk should be considered acceptable if there are no data indicating more than 50 % phytotoxic effect at the maximum application rate. If the results show more than 50 % effect for one species or clear indications of effects on more than one species, data requirements and assessment move to the next tier.

Tier 2: Quantitative risk assessment

This tier is a quantitative risk assessment following a TER approach. Both effects and exposure are expressed in terms of application rate (g/ha). Effects data are represented by ER50 values from the studies described under tier 2 in chapter 7.1, also expressed as g/ha. There are two options, a deterministic and a probabilistic approach, from which a choice should be made with regard to the data set (the probabilistic method is not always applicable).

Deterministic approach

If the TER based on the most sensitive species is greater than 5 then effects on non-target plants are considered acceptable. This trigger of 5 presupposes that at least 6 species have been tested. The trigger may be reduced if information on more species is available.

Probabilistic approach

Probabilistic methods that make use of the species sensitivity distribution would be straightforward in this assessment step as data from 6-10 species are available. Furthermore, a probabilistic approach is considered more suitable than the deterministic one to achieve the type of environmental goal mentioned above. This approach requires that log-normal or another defined type of distribution of the data has been shown to fit the data adequately. If the ED50 for less than 5 % of the species is below the highest predicted exposure level, the risk for terrestrial plants is assumed to be acceptable.

Tier 3: Higher tier risk assessment based on field studies

The third tier requires a higher tier risk characterisation and therefore, a case-by-case analysis. The ecological relevance of the observed effects, consequences on soil functions, and the potential for recovery are key elements for the assessment.

7.4 Risk mitigation options

In order to reduce exposure of non-target plants the options are similar to non-target arthropods in off-field areas:

- Buffer zones to sensitive areas
- Drift-reducing application techniques in the vicinity of sensitive areas.

As usual these measures are highly specific for Member State conditions.

8 Other non-target organisms

Effects on other non-target organisms (flora and fauna) believed to be at risk (Annex II 8.6)

There is a requirement for a summary of available data from preliminary tests used to assess the biological activity to be submitted. It is proposed that the summary should be presented in the monograph and any areas of concern highlighted. However, as non-target plants now are dealt with separately this summary in most cases will be very brief.

9 Terms and abbreviations

ECx	Effective concentration x % (concentration causing x % effect in a dose-response test); ECx is used as an overarching term referring to any kind of dose-response-modelling; ECx values may be specified with the first letter denoting the kind of endpoint (L = lethal), the second letter denoting the kind of exposure (C = concentration, D = dose, R = rate)
ED50	Effective dose 50 %
ESCORT	European Standard Characteristics of Beneficials Regulatory Testing
DT50, DT90	Disappearance time 50 % (90 %); the time it takes in a dissipation study until 50 % (90 %) of the initial amount or concentration has disappeared; the subscript f denotes field studies
f_{oc}	fraction of organic carbon
GAP	Good Agricultural Practice
HQ	Hazard quotient
IPM	Integrated Pest Management
LD50	Lethal dose 50 %
LR50	Lethal rate 50 %
MAF	Multiple application factor
NOEC	No observed effect concentration; highest concentration in a dose-response test which is not statistically different from the control
PEC	Predicted environmental concentration
PRA	Probabilistic Risk Assessment
TER	Toxicity/exposure ratio; subscripts denote time-scales (a = acute, st = short-term, lt = long-term)

10 References

Bakker FM, Feije R, Grove AJ, Jacobs G, Loose ED and van Stratum P (2002): A laboratory test to evaluate effects of plant protection products on mortality and reproduction of the predatory mite Hypoaspis aculeifer Canestrini (Acari: Laelapidae) in standard soil. J Soils Sediments, submitted.

Barrett KL, Grandy N, Harrison EG, Hassan S and Oomen P (eds) (1994): Guidance document on regulatory testing procedures for pesticides and non-target arthropods. From the ESCORT Workshop. SETAC-Europe, 51 p.

Bjørnlund L, Ekelund F, Christensen S, Suhr Jacobsen C, Krogh PH, Jagers op Aakkerhuis G and Johnsen K (2000): Interactions between saprotrophic fungi, bacteria and protozoa on decomposing wheat roots in soil influenced by the fungicide fenpropimorph (Corbel®): a field study. Soil Biol Biochem 32, 967-975.

Candolfi MP, Barrett KL, Campbell PJ, Forster R, Grandy N, Huet MC, Lewis G, Oomen PA, Schmuck R and Vogt H (eds) (2001): Guidance document on regulatory testing and risk assessment procedures for plant pretection products with non-target arthropods. From the ESCORT 2 workshop. SETAC, Pensacola, 46 p.

Chapman PM, Fairbrother A and Brown D (1998): A critical evaluation of safety (uncertainty) factors for ecological risk assessment. Environ Toxicol Chem 17, 99-108.

Connell DW and Markwell D (1990): Bioaccumulation in the soil to earthworm system. Chemosphere 20, 91-100.

CSTEE (2000): Scientific committee on toxicity, ecotoxicity, and the environment: Opinion on the available scientific approaches to assess the potential effects and risk of chemicals on terrestrial ecosystems. Opinion expressed at the 19th CSTEE plenary meeting - Brussels, 9 November 2000
http://europa.eu.int/comm/food/fs/sc/sct/outcome_en.html#opinions

Elkins NZ and Whitford WG (1982): The role of micro arthropods and nematodes in decomposition. Oecologia 55, 303-310.

EPFES (2002): Effects of plant protection products on functional endpoints in soils. Workshop held in Lisbon, 24-26 April 2002. (Proceedings in preparation.)

EPPO (2001): Guideline on test methods for evaluating the side-effects of plant protection products. No. PP 1/170(3). Honeybees. EPPO Bull 31, 323-330.

EPPO (2002a): Environmental risk assessment scheme for plant protection products. Chapter 8. Soil organisms and functions. EPPO Bull, in prep.

EPPO (2002b): Environmental risk assessment scheme for plant protection products. Chapter 11. Honeybees. EPPO Bull, in prep.

FOCUS (1996): Soil persistence models and EU registration. Final report of the work of the Soil Modelling Work group of FOCUS (FOrum for the Co-ordination of pesticide rate models and their USe). European Commision Document 7617/VI/96, 77 pp.

Ganzelmeier H, Rautmann D, Spangenberg R, Streloke M, Herrmann M, Wenzelburger H-J and Walter H-F (1995): Studies on the spray drift of plant protection products. Mitt Biol Bundesanst Land- Forstwirtsch Berlin-Dahlem, Vol 305.

Greig-Smith PW, Becker H, Edwards PJ and Heimbach F (Eds.) (1992): Ecotoxicology of earthworms. Intercept, Andover UK.

Hart ADM (ed.) (2001): Probabilistic risk assessment for pesticides in Europe: Implementation & research needs. Report of the European Workshop on Probabilistic Risk Assessment for the Environmental Impacts of Plant Protection Products (EUPRA). Published by Central Science Laboratory, Sand Hutton, York, UK, 109 pp.

Hart ADM and Thompson HM (1995): Significance of regurgitation in avian toxicity tests. Bull Environ Contam Toxicol 54, 789-796.

Jager T (1998): Mechanistic approach for estimating bioconcentration of organic chemicals in earthworms (Oligochaeta). Environ Toxicol Chem 17, 2080-2090.

Laskowski R (1995): Some good reasons to ban the use of NOEC, LOEC and related concepts in ecotoxicology. Oikos 73, 140-144.

Løkke H and Van Gestel CAM (1998): Handbook of soil invertebrate toxicity tests. Ecological and Environmental Toxicology Series, Wiley and Sons, Chichester, 281 pp.

Lynch MR (Ed) (1995): Procedures for assessing the environmental fate and ecotoxicity of pesticides. SETAC Europe, Brussels, 54 pp.

OECD (1998): Report of the OECD Workshop on Statistical Analysis of Aquatic Toxicity Data, OECD Series on Testing and Assessment, No. 10, 1998.

Oomen PA, De Ruijter A and van der Steen J (1992): Method for honeybee brood feeding test with insect growth-regulating insecticides. EPPO Bull. 22, 613-616.

Rautmann D, Streloke M and Winkler R (2001): New basic drift values in the authorisation procedure for plant protection products. In: Forster R, Streloke M (eds), Workshop on Risk Assessment and Risk Mitigation Measures in the Context of the Authorization of Plant Protection Products (WORMM). Mitt Biol Bundesanst Land-Forstwirtsch Berlin-Dahlem, 383, 133-141

Romijn CAFM.; Luttik R and Canton JH (1994): Presentation of a general algorithm to include effect assessment on secondary poisoning in the derivation of envrionmental quality criteria. 2. Terrestrial food chains. Ecotox Environ Safety 27, 107-127.

SCP (1999): Opinion of the Scientific Committee on Plants on endocrine disruption relevance in the context of Council Directive 91/414/EEC concerning the placing of plant protection products on the market. Opinion expressed by the Scientific Committee on Plants on 2 December 1999. http://europa.eu.int/comm/food/fs/sc/scp/out55_en.html

SCP (2002): Opinion of the Scientific Committee on Plants on the draft guidance document on risk assessment for birds and mammals under directive 91/414/EEC. Opinion adopted by the Scientific Committee on Plants, 24 April 2002. http://europa.eu.int/comm/food/fs/sc/scp/out125_ppp_en.pdf

Sheppard SC, Bembridge JD, Holmstrup M and Posthuma L (Eds.) (1997): Advances in earthworm ecotoxicology. SETAC Press

Tarazona JV, Hund K, Jager, T, S-Salonen M, Soares AMVM, Skaare JU andVigui M (2002): Standardizing chemical risk assessment, at last. Nature 415, 14.

Van Gestel CAM (1992): The influence of soil characteristics on the toxicity of chemicals for earthworms - a review. In: Becker H et al. (eds): Ecotoxicology of earthworms. Intercept Press, Andover.

Van Vliet PCJ, Gupta V and Abbott LK (2000): Soil biota and crop residue decomposition during summer and autumn in south-western Australia. Appl Soil Ecol 14, 111-124.